DOCUMENT RESUME

ED 368 908 CE 066 129

TITLE	Testing and Assessment in Vocational Education. Background Paper.
INSTITUTION	Congress of the U.S., Washington, D.C. Office of Technology Assessment.
REPORT NO	ISBN-0-16-043127-1; OTA-BP-SET-123
PUB DATE	Mar 94
NOTE	127p.
AVAILABLE FROM	New Orders, Superintendent of Documents, U.S. Government Printing Office, P.O. Box 37194, Pittsburgh, PA 15250-7954 (stock no. 052-003-01367-3: $7.50).
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC06 Plus Postage.
DESCRIPTORS	*Accountability; *Compliance (Legal); *Educational Change; Educational Legislation; Educational Policy; Evaluation Criteria; *Federal Legislation; Federal Regulation; Job Skills; *Measures (Individuals); Performance; Postsecondary Education; Public Policy; Secondary Education; State Government; *Vocational Education

ABSTRACT

This background paper provides a picture of general progress in vocational education assessment instruments and policies. Chapter 1 explains the paper's purposes and presents a summary of its contents. Chapter 2 traces the evolution of federal accountability requirements in vocational education law and explains the intent of the 1990 provisions. Chapter 3 profiles current state testing and assessment policies in vocational education. It analyzes how state assessment policies and practices are changing in response to Perkins Act requirements and considers how these changes could affect both other reforms in vocational education and the nature of learning and instruction in vocational education itself. Chapter 4 describes some testing and assessment resources available to measure various kinds of occupational skills, including tests developed by three main vendors. Chapter 5 explores alternative approaches for defining, teaching, and measuring broad technical skills and highlights issues that should be considered in moving toward assessments that meet needs identified by Congress. Chapter 6 describes how the U.S. Department of Education has implemented performance standards. One appendix lists legislative milestones related to accountability in federal vocational education programs. Another appendix explains acronyms. (YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 368 908

TESTING AND ASSESSMENT IN VOCATIONAL EDUCATION

BACKGROUND PAPER

CE 066 129

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

OFFICE OF TECHNOLOGY ASSESSMENT CONGRESS OF THE UNITED STATES

Technology Assessment Board of the 103d Congress

EDWARD M. KENNEDY, Massachusetts, *Chairman*
DON SUNDQUIST, Tennessee, *Vice Chairman*

SENATE	HOUSE
ERNEST F. HOLLINGS, South Carolina	GEORGE E. BROWN, JR., California
CLAIBORNE PELL, Rhode Island	JOHN D. DINGELL, Michigan
ORRIN G. HATCH, Utah	JIM McDERMOTT, Washington
CHARLES E. GRASSLEY, Iowa	AMO HOUGHTON, New York
DAVE DURENBERGER, Minnesota	MICHAEL G. OXLEY, Ohio

ROGER C. HERDMAN
(Nonvoting)

Technology Assessment Advisory Council

NEIL E. HARL, *Chairman*
Charles F. Curtiss Distinguished Professor
Iowa State University
Ames, Iowa

JAMES C. HUNT, *Vice Chairman*
Distinguished Professor, Health Sciences Center
University of Tennessee
Memphis, Tennessee

CHARLES A. BOWSHER
Comptroller General of the United States
Washington, D.C.

LEWIS M. BRANSCOMB
Director, Science, Technology & Public
Policy Program, Harvard University
Cambridge, Massachusetts

HERBERT (TED) DOAN
President (Ret.), The Dow Chemical Company
Midland, Michigan

JOSHUA LEDERBERG
Professor, Rockefeller University
New York, New York

MAX LENNON
President, Clemson University
Clemson, South Carolina

THOMAS J. PERKINS
General Partner
Kleiner, Perkins, Caufield and Byers
San Francisco, California

CHASE N. PETERSON
President Emeritus and Professor
University of Utah
Salt Lake City, Utah

WILLIAM H. ROBINSON
Acting Director, Congressional Research Service
The Library of Congress
Washington, D.C.

JOHN F.M. SIMS
Vice President, Marketing
Usibelli Coal Mine, Inc.
Fairbanks, Alaska

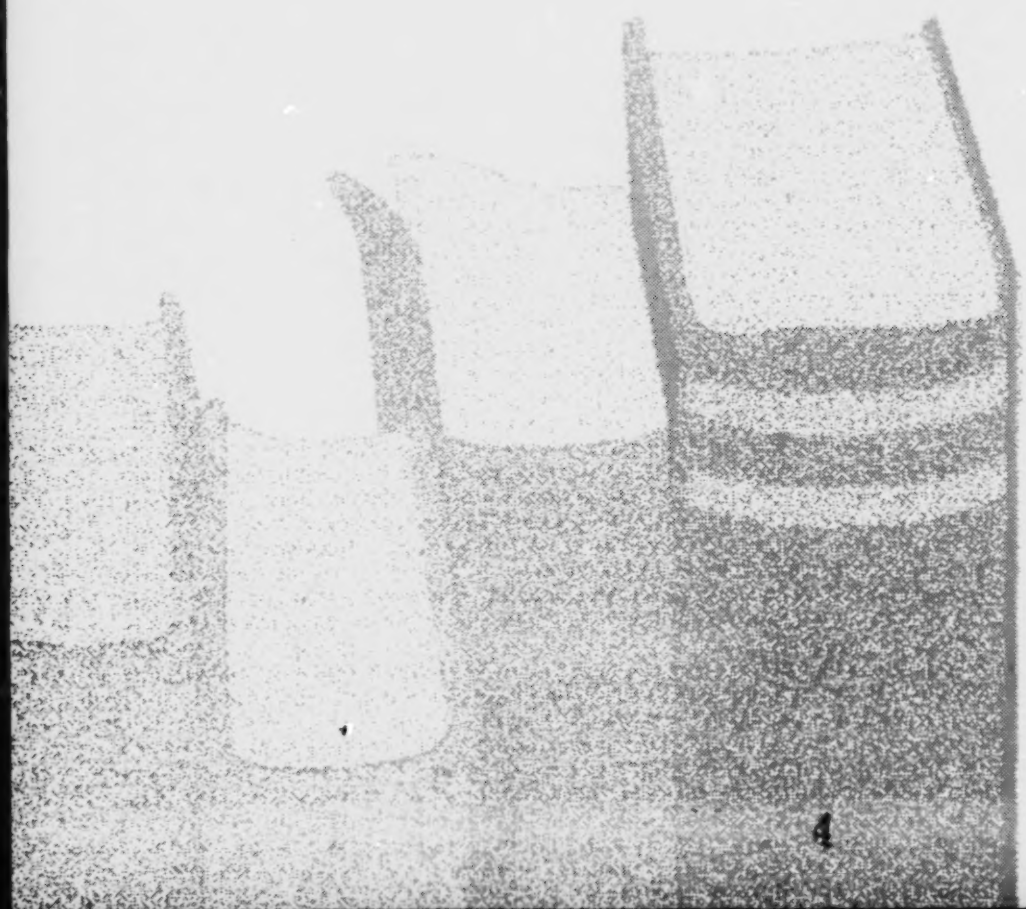
MARINA v.N. WHITMAN
Professor, Institute of Public Policy Studies
University of Michigan
Ann Arbor, Michigan

The views expressed in this background paper are not necessarily those of the Board,
OTA Advisory Council or individual members thereof.

BEST COPY AVAILABLE

TESTING AND ASSESSMENT IN VOCATIONAL EDUCATION

BACKGROUND PAPER



OFFICE OF TECHNOLOGY ASSESSMENT CONGRESS OF THE UNITED STATES

Recommended Citation:
U.S. Congress, Office of Technology Assessment, *Testing and Assessment in Vocational Education*, OTA-BP-SET-123 (Washington, DC: U.S. Government Printing Office, March 1994).

For sale by the U.S. Government Printing Office
Superintendent of Documents, Mail Stop: SSOP, Washington, DC 20402-9328
ISBN 0-16-043127-1

Foreword

The world of work is changing. As this occurs, educators must rethink how to prepare young people for employment. For young people who do not plan to obtain a college degree, it is no longer clear what type of employment will provide satisfying work that can lead to financial independence. And it is perhaps even less clear how students can be prepared for a work environment characterized by change.

As part of its growing concern for noncollege bound youth, Congress has begun to revisit and revise legislation that supports vocational education. In the 1992 revision of the Perkins Act, the House Committee on Education and Labor asked that OTA provide information on the types of tests and assessment instruments used to define the outcomes and effectiveness of these programs. By looking at these instruments, Congress can obtain insight. What we test is what we teach.

Accordingly, OTA has compiled this background paper on tests and assessments for secondary vocational education. It includes a state-by-state survey of assessment instruments, as well as information on new instruments under development by test vendors. Additionally, OTA has reviewed the emerging theories attempting to define "broad technical skills."

The survey data suggest some shifting away from the traditional, performance-based measurements of vocational education. This is occurring at a time when performance-based instruments are being introduced in other school subjects, in order to give a more useful and accurate indication of student achievement. OTA also found that there is no consensus on the meaning of "broad technical skills," and that a number of approaches are being used to try and prepare vocational education students for a changing workplace.

This background paper should help to inform the debate on vocational education and the school to work transition. These topics will be of substantial importance to Congress and the nation in the coming years, and OTA is glad to contribute to this discussion.

Roger C. Herdman
Director

Project Staff

John Andelin[1]
Assistant Director
Science, Information, and Natural Resources Division

Peter Blair[2]
Assistant Director
Industry, Commerce, and International Security Division Division

Nancy Carson
Program Manager
Science, Education, and Transportation

CONTRACTORS

Richard Kimbell
Goldsmiths' College
University of London

Nancy Kober
Charlottesville, VA

PRINCIPAL STAFF

JOHN WIRT
Project Director

OTHER CONTRIBUTING STAFF

Elizabeth Eisner
Detailee, U.S. Department of Education

C. Maxine Williams
In-house contractor

ADMINISTRATIVE STAFF

Marsha Fenn
Technical Editor

Tamara Kowalski
Research Secretary

Gay Jackson
PC Specialist

[1]Through August 1993.
[2]After August 1993.

Workshop Participants and Reviewers

WORKSHOP ON CAPABILITY FOR USING TECHNOLOGY, MARCH 9, 1993

Sherrie Gott, *Chairperson*
Air Force Human Resources Institute

Henry Braun
Educational Testing Service

Larry Hirschorn
Center for Applied Science

Richard Kimbell
University of London

Roger Shank
Northwestern University

Joyce Shields
Hay Systems, Inc.

Ronald Todd
Drexel University

Shoshana Zuboff
Harvard Business School

REVIEWERS

Barbara Border
Educational Leadership Consultants

Neal B. Carey
Center for Naval Analysis

Cheryl Fields Tyler
American Electronics Association

Brenda Hattaway
Vocational-Technical Education Consortium of the States

Gary Hoachlander
MPR Associates

Richard Jones
New York State Department of Education

Darrell Parks
Ohio State Department of Education

Evelyn Wells
National Occupational Testing Institute

Joel West
American College Testing

Lauress Wise
Manpower Data Systems Center

NOTE: OTA appreciates and is grateful for the valuable assistance and thoughtful critiques provided by the workshop participants, reviewers, and contributors. The participants, reviewers, and contributors do not, however, necessarily approve, disapprove, or endorse this report. OTA assumes full responsibility for the report and the accuracy of its contents.

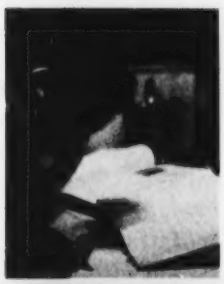
Contents

1 Purposes of the Background Paper 2
- Evolution of Performance-Based Accountability in Federal Vocational Education Law 4
- Testing and Assessment Resources in Vocational Education 6
- Current State Assessment Policies in Vocational Education 11
- Emerging State Responses to Perkins Performance Standards 14
- Defining and Assessing Broad Technical Skills 17

2 The Evolution of Performance-Based Accountability Requirements in Federal Vocational Education Law 21
- The Current Requirements 22
- Legislative History 26
- Changing Themes 32

3 State Policies on Testing and Assessment 35
- Design and Methodology of the State Survey 37
- Origins of Testing and Assessment Policy in Vocational Education 38
- Differences Among the States in Policy on Testing and Assessment 41
- State Resources for Testing and Assessment 42
- What Skills are Assessed? 44
- Where in the Curriculum are the Skills Assessed? 48
- Why are the Skills Assessed? 51
- What Form of Measurement is Used? 53
- How are Practices of Testing and Assessment Changing? 55
- What Skills Will Be Assessed? 59

4 Case Studies: Three Vendors of Tests for Occupational Competencies 67
- Work Keys 67
- Vocational-Technical Education Consortium of the States 72
- National Occupational Competency Testing Institute 76
- Findings 80



5 Broad Technical Skills 83
Vocational Aptitudes 84
Occupational Maps 91
Core Occupational Skills 94
Design and Technology 95
Cognitive Skills 97
Final Note 101

6 Implementation of Performance Standards and Measures 103
Technical Assistance Efforts 104
ED Regulations and Monitoring 105
Research and Demonstration 107
Business and Education Standards 107
Goals 2000: Educate America Act 108
Conclusion 109

APPENDICES

A Legislative Milestones Influencing Accountability in Federal Vocational Education Programs 111

B List of Acronyms 114

Photographs supplied by the American Vocational Association and Future Business Leaders of America.

Summary 1

In 1990, Congress enacted amendments to the Carl D. Perkins Act requiring states to measure the effectiveness of vocational education and to set performance standards for vocational programs. The law requires using these standards to evaluate the effectiveness of local vocational education programs and stimulate "program improvement" efforts where deficiencies are found.

The requirements for standards and measures are part of a larger thrust in the 1990 amendments to strengthen accountability for vocational programs. Other reforms advocated in the legislation include integrating academics and vocational education, and creating "tech prep" programs. At the high school level, both involve reorganizing vocational programs around broader sets of academic and technical skills leading to advanced training and highly skilled jobs.

The requirements for standards and measures are likely to have a significant impact on the amount and type of *testing and assessment* occurring in vocational education. High-quality assessments are crucial; incomplete or inaccurate assessment results could lead to unfair or misleading conclusions about which local programs need improvement. And depending on which type of tests are used and how they are applied, the implementation of performance standards could either promote or impede the other reform goals of the 1990 law.

To help Congress determine whether better assessment instruments should be and can be developed for vocational education, section 423 of the Perkins Act directed the Office of Technology Assessment (OTA) to study tests used in vocational programs, particularly those designed to measure broad technical

skills. Congress asked OTA to analyze the uses of these tests, identify trends in vocational assessment, and identify policy issues relevant to improving test development and quality.

PURPOSES OF THE BACKGROUND PAPER

This background paper responds to the congressional request by providing a picture of general progress in vocational education assessment instruments and policies. The paper has six major purposes, each explored in greater detail in a subsequent chapter:

- to trace the evolution of federal accountability requirements in vocational education law and to explain the intent of the 1990 provisions (ch. 2);
- to profile current state testing and assessment policies in vocational education (ch. 3);
- to analyze how state assessment policies and practices are changing in response to Perkins Act requirements and to consider how these changes could affect both other reforms in vocational education and the nature of learning and instruction in vocational education itself (ch. 3);
- to describe some of the testing and assessment resources available to measure various kinds of occupational skills, including tests developed by three main vendors (ch. 4);
- to explore alternative approaches for defining, teaching, and measuring broad technical skills and to highlight issues that should be considered in moving toward assessments that meet needs identified by Congress (ch. 5); and
- to describe how the Department of Education has implemented performance standards (ch. 6).

In the course of describing this progress, the paper also raises a number of broader issues of how the policies of testing and assessment adopted by states and localities may turn out to affect the implementation of performance standards, accountability in the federal legislation, and ultimately the nature and content of vocational education. Understanding these relationships is particularly important, because the purpose of the standards is to make decisions about the effectiveness of programs at the state and local levels.

It is also important to note that testing and assessment is conducted in vocational education for many purposes. These include instructional diagnosis, grading student performance, certification of student competence for employers, and various other policy purposes at the state and local levels. OTA's survey covered all practices of testing and assessment affected by state policy; not just testing and assessment tied to performance standards.

Defining Types of Skills

Several kinds of skills can be assessed in vocational programs, reflecting the multiple goals of vocational education. For purposes of this paper, OTA has grouped them into four types (see box 1-A). The first type are *academic skills*, primarily the areas of reading, writing, and mathematics. The other three types are various kinds of occupational skills: *vocational skills*, which tend to be job specific; *generic workplace skills*, which encompass employability skills such as positive work attitudes, as well as teamwork, effective communication, and other kinds of general workplace competencies; and *broad technical skills*, which are the core skills and understandings of technology, information, and organization needed to perform effectively within an industry or range of occupations.

Methodology and Limitations of This Study

This background paper draws on several resources. Although assessment practices in vocational education have not been the subject of much prior research, OTA reviewed the available studies. Of particular relevance were a 1992 study of state implementation of performance standards

Box 1-A—Types of Skills Assessed in Vocational Education

Academic Skills

The knowledge and skills in the traditional academic subject areas of English, mathematics, science, history, and so forth that students are expected to acquire in school. In vocational programs, these skills may be tested for or assessed directly using standard methods or in the context of occupational problems or situations.

Occupational Skills

As used in this report, occupational skills refer to knowledge and skills other than academic needed to perform effectively in jobs; in other words, either vocational, generic workplace, or broad technical skills

Vocational Skills—The specific knowledge and skills required in the performance of particular jobs or groups of jobs within a certain occupational area.

Generic Workplace Skills—There are two types: employability skills, such as positive work attitudes and knowledge of how to find a job; and general competencies, such as ability to work in teams, communicate well with others, and solve problems. One definition of these general competencies is the five workplace competencies of the Secretary's Commission on Achieving Necessary Skills (SCANS).[1]

Broad Technical Skills—The core skills and understandings of technology, information, organization, and economics needed to perform effectively within an industry or group of industries.

[1] The five workplace competencies identified by SCANS are that effective workers can productively use resources, interpersonal skills, information, systems, and technology. Secretary's Commission on Achieving Necessary Skills, *What Work Requires of Schools* (Washington, DC: U.S. Government Printing Office, June 1991).

by MPR Associates;[1] a survey of state competency testing in vocational education by Mississippi State University;[2] and a survey of state initiatives for industry skill standards by the National Association of State Directors of Vocational Education.[3]

To obtain more information on current state testing practices and proposed changes, OTA conducted a survey of the 50 states and the District of Columbia, following up with telephone interviews. OTA also conducted case studies of three major test vendors in vocational education.

This paper is exploratory, intended to provide basic descriptive information and point to directions for future policy analyses. Many of the issues raised are complex and will require further study. For example, data on test use are not available for special populations of students, an area of keen congressional interest. More extensive study of local assessment practices in vocational education would also be beneficial. Because states vary greatly in their influence on local assessment practices, state surveys can provide only a partial view of local practices.

Over time, the testing and assessment approaches and instruments chosen by states and localities will vary. Careful attention to the effects of these choices on the implementation of performance standards and on the nature of curricu-

[1] Mikala L. Rahn et al., MPR Associates, Inc., "State Systems for Accountability in Vocational Education," prepared for the U.S. Department of Education, Office of Vocational and Adult Education, December 1992.

[2] Rebecca Love-Wilkes and Ronda Cummings, Research and Curriculum Unit for Vocational, Technical, and Adult Education, Mississippi State University, "1990 State of the Art Report on Statewide Student Competency Testing in Vocational and Technical Education," prepared for the National Network for Curriculum Coordination in Vocational-Technical Education and the Southeast Curriculum Coordination Center, October 1990.

[3] Barbara Border, *Education-Driven Skill Standards Systems in the United States*, prepared for the U.S. Department of Education (Washington, DC: National Vocational Technical Education Foundation, October 1993).

lum and instruction in vocational education will be required to determine whether the legislative goals have been reached.

EVOLUTION OF PERFORMANCE-BASED ACCOUNTABILITY IN FEDERAL VOCATIONAL EDUCATION LAW

The current requirements for performance standards and outcome measures are the most recent stage in the evolution of accountability requirements in the federal vocational education law. **Every major vocational education law since the Smith-Hughes Act of 1917 has included accountability requirements, but the tools for achieving accountability have become more outcome-oriented over time.**

In the early years of federal support, when the primary goal was to encourage growth in vocational education, accountability was enforced by regulating program "inputs." States were directed to establish and adhere to minimum requirements for vocational teacher qualifications, classroom equipment, and instructional hours per week. These requirements gave federal program administrators a tool that they wielded aggressively to shape the growth of vocational education, thereby helping to establish the concept of federal funds as carrot and stick.

The first attempt to define and look systematically at outcomes of vocational education occurred with the Vocational Education Act of 1963, which introduced a requirement for periodic evaluation of local programs. Program offerings were to be reviewed in terms of "... current and projected manpower needs and job opportunities . . ."[4] in relevant occupational fields; however, the law did not specify how these reviews were to be conducted. The 1968 amendments took another step toward defining outcomes by limiting federal support to local programs that could be demonstrated "... to prepare students for employment . . . or be of significant assistance [to students] in making an informed and meaningful occupational choice."[5] In this way the amendments emphasized that the primary purpose of vocational programs was to provide students with the specific skills needed for real jobs, not just with general learning in the manual arts.

The 1976 law further sharpened the focus on outcomes by specifying that the mandated local program reviews should examine the extent to which program completers and leavers: a) found employment in occupations related to their training, and b) were considered by employers to be well trained.

The Perkins Act of 1984 explicitly directed states to develop measures of program effectiveness, such as the occupations to be trained for, the levels of skills to be achieved, and the "... basic employment competencies to be used in performance outcomes, which will reflect the hiring needs of employers . . ." (section 113). Foreshadowing the current movement to define skill standards for various industries, the 1984 law also required states to establish technical committees of business and industry representatives; these committees were to develop inventories of skills for "priority occupational fields," which could be used to "define model curricula."

By 1990, Congress had concluded that prior calls for change had not spurred significant improvements in the quality of vocational education. Influenced by experiences with outcome-based accountability in other federal education and training programs, Congress amended the Perkins Act to require states, within 1 year, to develop and implement statewide systems of "core standards and measures" that defined the student outcomes expected in local programs.

In delineating the types of outcomes that states could select, Congress endorsed a broad view of the purposes of vocational education that encom-

[4] 77 Stat. 406.

[5] 82 Stat. 1076.

passed academic achievement, dropout prevention, and preparation for higher education, as well as job preparation. Thus, the law states that the standards must include a minimum of two outcome measures: 1) a measure of learning and competency gains, including student progress in basic and more advanced academic skills; and 2) a measure of one or more of the following—job or work skill attainment or enhancement, school retention or secondary school completion, placement into a job, additional training or education, or military service. State systems also are required to include incentives or adjustments that encourage service to special populations, although the legislation offers no guidelines on how this should be done.

By including academic outcomes in the performance standards and placing priority in other parts of the legislation on integrating academic and vocational education, Congress signaled a major, new direction of federal policy on vocational education. The intent of the policy is that students who take vocational education should have the same academic knowledge and skills as other students. It is a statement that students who graduate from vocational programs should be as well equipped for their future lives of work and learning as other students.

The 1990 legislation marks a significant turning point in federal accountability by explicitly tying the process of state and local review to standards based on outcomes. Beginning after 1993-94, each local recipient of Perkins Act basic grant funding must use the statewide standards and measures to evaluate annually the effectiveness of its vocational programs. (Local recipients may use federal funds, to a reasonable extent, to conduct these reviews.) Eventually, local recipients who are not making substantial progress toward the standards must develop a "program improvement" plan identifying changes to be made in the following year.

The requirement for standards is also significant as much for what it does *not* require as for what it does. First, Congress did not authorize the Secretary of Education to issue national standards and measures, but instead gave states considerable flexibility to design their own systems and select from a range of outcomes. Only two outcome measures are required. (In practice, however, most states have adopted multiple standards and measures.)

Second, the main purpose of the performance standards is to make decisions about programs. The results of the performance standards are specifically tied to the annual review requirement, which has been in the federal legislation for some time. The standards are not intended to certify or credential individuals. (This is in contrast to current proposals for industry skill standards, which would be used for individual credentialing.) Although testing of students will be necessary to satisfy some of the performance standards developed by the states, the law makes clear that the results are to be used primarily to evaluate and improve local programs. It is up to the state to decide whether to implement performance standards as a system of student certification or to incorporate student certification functions into their overall plan. According to the OTA state survey, at least two states, Pennsylvania and New Jersey, have chosen to implement performance standards as systems of student certification.

Third, Congress chose not to link the vocational education performance standards to federal funding or any other incentives or sanctions. As House and Senate reports make clear, no authority exists for states ". . . to apply sanctions in connection with the utilization of measures and standards at the local level."[6] The mild consequence that was attached to the local program evaluations—state intervention through joint state-local program improvement plans—was not intended ". . . to be punitive in nature, but rather to

[6] U.S. Congress, House Committee on Education and Labor, "Applied Technology Education Amendments of 1989," H. Rept. 101-41, Apr. 28, 1989, p. 4.

encourage an infusion of resources from the state [for] programs that are in need of assistance and improvement."[7]

Fourth, the legislation is not intended to position Congress as the distant, final judge of local processes. Thus, the local level, not the federal level, is the primary arena for conducting evaluations, reviewing evaluation data, and carrying out program improvement (with the involvement of the state if needed). The act does not require recipients to submit the results of evaluations to the Secretary, nor does it direct evaluation results to be structured so as to yield a national database on program effectiveness or a national system of performance standards. National information needs are to be met though other mandated studies and research.

In passing the legislation, there was some support in Congress for encouraging or even requiring the development of a national system of performance standards for vocational education based on the performance standards to be developed by the states. Congress decided to provide states with a great deal of flexibility in defining their performance standards and not require the development of a national system. The issue was resolved by including in the final legislation a study to evaluate the quality and comparability across states of the performance standards and measures adopted by the states. The presence of the study suggests that, in the future, consideration could be given to forming a national system of performance standards for vocational education. In considering such a step, two of the important criteria could well be the extent of agreement among the states on vocational outcomes and their capabilities for measuring those outcomes.

Both the flexibility given the states and the possibility of expanding the performance standards into a national system lead to a number of important future policy questions. States will have to make difficult decisions on outcomes with a great deal of latitude and not much experience. It is an open question whether information from student testing and assessment will prove to be useful for making decisions about improvements in local programs. It is also an open question whether the performance standards and measures developed for local program improvement could or should be used to develop a common core of competencies and indicators at the national level. Finally, it is an open question what the effects on the nature and content of vocational education may be from using testing and assessment information for purposes of accountability, especially given the imperfect quality of the available methods and instruments of testing and assessment. At this point, the effects are impossible to predict.

TESTING AND ASSESSMENT RESOURCES IN VOCATIONAL EDUCATION

The federal legislation is silent about the types of testing and assessment resources needed to measure student outcomes and implement performance standards. Nevertheless, the new requirements will place substantial burdens on state and local testing and assessment instruments, programs, and practices. It seems clear that the capacity for testing and assessment must increase in response, but by how much and in what direction is uncertain. Because there has been little systematic research on vocational education testing and assessment issues, it is not even clear what resources currently exist, what the range of testing practices is, and how these practices compare with the rest of education. As a starting point, OTA pulled together existing information on testing resources in vocational education and conducted a survey of state practices.

[7] U. S. Congress, Senate Committee on Labor and Human Resources, "Carl D. Perkins Vocational Education Act Amendments of 1989," S. Rept. 101-221, Nov. 21, 1989, pp. 22-23.

Overview of Testing in Vocational Education

After reviewing evidence from its state survey and other research, OTA concludes that **testing and assessment practices in secondary vocational education differ considerably from the rest of education in their features, origins, and applications. In fact, the best assessment practices in vocational education resemble the alternative forms of assessment just now being explored for the rest of education.** However, the quality of these assessment practices varies greatly among states and localities.

Like the rest of education, vocational education programs use both written, short-answer forms of testing and diverse methods of performance assessment, but vocational education relies much less heavily on short-answer methods of testing and more on assessment. The answer formats of the written tests are typically matching or multiple choice. A growing amount of this *written testing*, as it will be called, is done using either instruments that are centrally developed by states or test vendors, or locally adapted forms of those instruments (see box 1-B). The centrally developed instruments are produced through iterative cycles of writing and revision, but the resulting instruments are typically much less standardized and easier to adapt than are the highly standardized and secure tests that are so common in academic education. Further complicating the picture, most of these centrally developed instruments include both written and performance exercises; however, when they are used, the written portions generally predominate. At the local level, teachers may also prepare and use their own written tests. The centrally developed written tests and the adapted versions of them produced at the state and local levels are the focus of this report rather than locally produced written tests.

The diversity of assessment methods utilized is broad. The range includes the preparation of student profiles and portfolios, structured teacher ratings of student capabilities demonstrated in the course of regular classroom work, evaluated student projects, and even organized competitive events.

In contrast to the rest of education, both the written testing and the diverse forms of assessment used in vocational education are nearly all *criterion-referenced* rather than *norm-referenced*—meaning that they are designed to measure whether students have the knowledge and skills needed for particular jobs rather than how they perform relative to other students.

The testing and assessment done in vocational education stems from very different origins than testing and assessment in academic education. The roots of standardized academic testing lie in the mental testing movement in psychology and education.[8] The source of testing and assessment in vocational programs is the *competency-based* movement in vocational training. In properly conducted competency-based vocational education, the curriculum content, test items, and performance exercises are derived from analyses of actual tasks performed by people in specific jobs or occupational areas. Ideally, there is a very close alignment of instruction, assessment, and job tasks. As a result, teachers have a much more central role in judging student performance than in standardized academic testing.

In the best competency-based programs, it is skills learned, not time spent, that drives the pace of instruction for individual students. In this respect, the philosophy of competency-based instruction and assessment is wholly different from the philosophy of whole-class instruction with mass testing at fixed points in the curriculum. **In vocational education, testing and assessment are not after the fact, external processes of inspection but integral parts of the**

[8] U.S. Congress, Office of Technology Assessment, *Testing in American Schools: Asking the Right Questions*, OTA-SET-519 (Washington, DC: U.S. Government Printing Office, February 1992), ch. 4.

Box 1-B—Glossary of Testing Terms

Competency Testing

Competency testing is the administration of written tests to determine whether students have the knowledge and skills required to perform effectively in a job or occupational area. The individual items of the test are derived from analyses of the specific tasks involved in performing those jobs. The answer format of these tests is typically *closed ended,* that is, multiple choice or matching. The results of the testing may be used for purposes of improving instruction or instructional programs, documenting or reporting student achievement, or certifying student capabilities to employers.

Competency Assessment

This type of assessment uses one or more methods of observation, rating, and recording of actual performances to determine the capabilities of students for performing well in a job or occupational area and conveying the results to others. The performances observed and rated may be part of the student's regularly assigned classroom or project work, an organized event of some kind, or occur in response to a problem or task situation especially assigned for the purpose of assessment. Student performances are typically rated or evaluated according to a structure of valued competencies derived from the analyses of the tasks actually performed on the job and a scale of performance levels. The preparation of student profiles, portfolios, or other forms of documenting the results of the assessment and explaining the students' performances may be part of the assessment process. The results of the assessment may be used for purposes of improving instruction or instructional programs, documenting or reporting student achievement, or certifying student capabilities to employers.

Academic Testing

In academic testing, written instruments are used to measure the knowledge and skills of students in the traditional academic subject areas of mathematics, writing, science, literature, history, and so forth. The answer format of these tests is typically multiple choice, matching, or some other such closed-ended form of response.

process of education—a goal only now being advanced in academic education.

In the competency-based tradition, assessment includes carefully designed performance exercises or tasks, instructor or juried assessment of completed student projects, and teacher assessment of regular classroom work using systematic methods of rating. There is also a strong tradition of organized events in which students compete for recognition and reward. The competency-based model attempts to systematize these various forms of performance assessment by providing the instructor or judge with performance scales and lists of valued competencies that can be "checked off" or rated as students perform various tasks on demand or over a period of time.

For a number of years, interest has also been growing in vocational education in the idea of providing students with profiles of their competencies and encouraging them to build portfolios of their accomplishments to use in job seeking.

Conclusions cannot yet be reached about whether these various methods of performance assessment used in vocational education are more or less reliable or valid than the written testing that is done. Before conclusions can be drawn about which methods are best for which purposes, closer investigations must be conducted in vocational education of the consistency and relevance of different assessment methods, and their actual applications in vocational education.

Standardized Testing

Standardized tests use written instruments that are developed according to principles of test theory and administered according to uniform procedures to provide comparability of results among test takers. Standardized testing is principally used to measure academic achievement.

Performance Assessment

In performance assessment, student performances are rated in response to an "on-demand" task; that is, all students respond to the same task (or prompt) that has been given to them expressly for the purpose of observing and evaluating their capabilities for performance in certain areas. The tasks are developed through iterative cycles of trial and revision. Performance assessments may be academically or occupationally oriented. The tasks imposed are generally situated to simulate real environments and are *open ended* in that students may respond to them in many different ways. They also typically involve more complex levels of thinking and doing than are possible with closed-ended testing.

Criterion-Referenced Tests or Assessments

These tests or assessments focus on "... what test takers can do and what they know, not how they compare to others." Criterion-referenced tests or assessments are designed to show how a student is doing relative to competencies required on the job or specified educational goals or objectives.[1]

Norm-Referenced Tests

These tests are designed to compare one student's performance with the performances of a large group of students. Norm-referenced tests are developed to make fine distinctions between students' performances and accurately pinpoint where a student stands in relation to a large groups of students. Assessment is almost by definition not norm referenced.[2]

---

[1] U.S. Congress, Office of Technology Assessment, *Testing in America's Schools: Asking the Right Questions*, OTA-SET-519 (Washington, DC: U.S. Government Printing Office, February 1992), pp. 169-170.

[2] Ibid., pp. 168-169.

The critical issues in *performance assessment* are the comparability of judgments from instructor to instructor and program to program, and the correspondence of those judgments with any standards that have been set. In some cases, business representatives or parents may be involved. With sufficient training for judges, group methods of judging, and statistical checks on the consistency of ratings, it is possible to achieve satisfactory levels of consistency in rater judgment across units in performance assessment.

The critical issues in *written testing* are the relevance of test items to capabilities for actual job performance and the long-term effects of the testing method on teaching and learning in vocational programs. Written test formats are generally thought to be best for measuring factual knowledge and certain forms of cognition and reasoning, which may or may not be closely related to "know-how" and capabilities needed for complex and extended performance in the workplace (and in life).

Resources From Test Vendors

Test vendors supply some of the testing and assessment materials used in vocational education. Three of the best known testing organizations in vocational education are the Vocational-Technical Consortium of the States (V-TECS), the National Occupational Competency Testing Institute (NOCTI), and American College Testing (ACT). NOCTI produces the Student Occupa-

tional Competency Achievement Testing (SOCAT) materials. ACT is in the early stages of marketing a new testing program called Work Keys, which measures general workplace competencies. The testing and assessment products of these three organizations are distinctly different.

V-TECS and NOCTI utilize structured methods of job analysis to develop competency-based materials for vocational education assessment. Both organizations produce performance exercises as well as short-answer written test items, although more extensive resources are available for written testing.

NOCTI is the more conventional of the two. The tests it develops and sells to states and local programs are secure, consisting of fixed sets of items derived from job analyses in relevant fields. The SOCAT tests produced by NOCTI are available for 71 specific occupations.

V-TECS, by contrast, does not sell "tests," but instead provides its 23 state members and constituent local programs with 35 different "test item banks," which they may use to construct tests reflecting their own state or local priorities. (V-TECS testing materials are also available for purchase by nonmembers.) Each V-TECS item bank is specific to a job area or occupation. Items are scrutinized for occupational relevance and possible bias and pilot-tested for consistency of response. The initial development of V-TECS materials is done by vocational education agencies in member states.

V-TECS also makes available lists of competencies and performance standards by occupation. These V-TECS catalogs, as they are called, are available for over 200 occupations.

In the OTA state survey, state personnel frequently reported devoting substantial efforts to adapting, redeveloping, and expanding V-TECS catalogs and item banks, or using them in conjunction with competency lists, tests, or items from other sources. The most common reason given for doing so is that neither the V-TECS materials nor those from other sources adequately reflect state and local priorities among different areas of knowledge and skills. Whether this reinvention and adaptation is genuinely useful or merely duplicative is impossible to say from the data available. Local priorities undoubtedly differ from state and national ones. Moreover, several studies have found that the process of reinvention is essential to the thorough implementation of innovations—"to understand is to invent." Still, questions remain about whether this reinvention affects the comparability of assessment results from place to place and how much of it is really necessary.

The new battery of Work Keys tests being developed by ACT differs from the V-TECS materials in some important respects. The Work Keys tests generally fit the basic model of "written testing," because of the thoroughness with which they are being developed, their requirements for standardized administration, and their centralized scoring and secure nature.[9] But several of the Work Keys tests involve innovative methods of response, such as listening to audiotapes and transcribing what is heard, viewing videos to provide a context for extended forms of multiple-choice questions, and watching demonstrations to assess learning from observation. The Work Keys system is just now being implemented in several states and local programs.

The main innovation of Work Keys is its focus on general workplace competencies, such as "applied technology" and "teamwork," rather than on job-specific skills. ACT plans to provide a means to compare profiles of skills needed in different job areas with profiles of knowledge and skills demonstrated by test takers. In short, Work Keys uses a different approach from V-TECS and SOCAT to link test content with job skill requirements.

How much influence do these three major vendors have on testing and assessment practices in vocational education? Available evi-

[9] ACT plans to offer local scoring of the Work Keys tests as an option in the future.

dence suggests that their impact is limited so far. V-TECS *currently* appears to be having the greatest influence through its deliberate strategy of modeling good competency testing practices and providing resources that states and local programs can use to develop their own assessment programs. Not all states belong to V-TECS, however, and V-TECS has tests for only 35 of the more than 200 occupational areas in which job competencies have been defined.

The most concrete estimates of the number of students taking vendor tests are available for the SOCAT tests, which are returned to NOCTI for scoring. Although NOCTI has many other clients for its testing products, the number of SOCAT test takers in schools is not large. For 1992, NOCTI reports that 9,015 secondary and postsecondary students took SOCATs; in that same year about 720,000 high school seniors were vocational students.[10]

Work Keys is too new to say how extensive its impact will be, but at least two states, Ohio and Tennessee, have adopted portions of it for statewide use, and many more are considering it.

CURRENT STATE ASSESSMENT POLICIES IN VOCATIONAL EDUCATION

State vocational education agencies are another source of testing and assessment resources for vocational programs. Through its survey and other studies, OTA has collected basic descriptive information about state assessment policies for vocational education as another means of determining the resources available for vocational education assessment.

Based on its survey and other evidence, OTA finds that state testing policies for vocational education are quite different from state testing policies for elementary and secondary education in general.

At the elementary and secondary levels, state Departments of Education commonly fund and operate programs for mass testing of students at various grade levels in various subjects. The purpose of much of this testing is to demonstrate accountability, and results are reported to the public.

By contrast, no state vocational education agencies directly administer a program of mass testing or assessment of all students at a fixed point in time. In most states, the primary assessment responsibility of the agency is to set policies for local programs to follow. Most state agencies also provide assessment resources to local programs, such as competency lists, test item banks, and tests with instructional materials. The main purposes of these state policies are to evaluate programs and courses, and assess and certify student progress—not to demonstrate accountability.

Categories of State Policies

Responses to the OTA survey reveal a variety of testing instruments and policies among the states. The 50 states and the District of Columbia reported a total of 92 different components of testing or assessment for academic skills and the 3 different kinds of occupational skills in their state programs, or an average of about 2 per state. Generally one of these components is for academic skills and the other is for occupational skills. Some states have more than one component of testing or assessment in each of these areas.

OTA finds that state assessment policies can be grouped into four distinct categories:

1. Eighteen states *mandated or strongly encouraged written forms of competency testing* for occupational skills in local programs in the 1992-93 school year.
   - All of these states favor written methods of testing over alternative forms of assessment for occupational skills.
   - All provide local programs with competency tests or access to a competency test item bank.

[10] Vocational students are defined as students who took more than four credits of vocational education in their high school careers.

- These states enroll about one-half of all high school students.
- Approximately one-half of them are known as leaders in the development of competency-based testing in vocational education.
- The sophistication and comprehensiveness of the testing programs in these 18 states varies greatly. Some have comprehensive programs of testing and assessment consisting of three or four components, such as a competency-based test item bank and supporting resources for over 50 occupations; student profiles or portfolios; and a strategy for obtaining the test scores of vocational students from the statewide academic testing program. Other states in the category may only offer local programs a single written test of employability skills.

2. Fifteen states *mandate assessment* of occupational skills in local programs without specifying how this should be done.
   - These states tend to encourage a diversity of approaches to assessment without favoring some methods over others. Various forms of *assessment*, rather than written testing, are generally encouraged.
   - All of these states require assessment of the occupational skills of students, but allow local programs to choose their own method or methods.
   - These states are much less likely to provide local programs with competency tests or item banks than states in category 1, and are more likely to provide materials for developing student profiles or portfolios.
   - These states enroll about one-quarter of all high school students.
3. Ten states *encourage assessment* of occupational skills in local programs without specifying how this should be done.
   - Like the previous category, these states encourage diverse approaches to assessment rather than written testing.
   - Testing and assessment are only encouraged, not required. The encouragement given to testing and assessment is generally not strong.
   - The only assessment resource that these states are likely to provide local programs is testing or assessment materials that come with instructional resources for a competency-based curriculum.
   - These states enroll about one-eighth of all high school students.
4. Eight states had *no specific policy or program* in school year 1992-93 to encourage or require testing or assessment of occupational skills.
   - In 1992-93, these states provided no resources to local programs for assessing occupational skills.
   - Most of these states have established performance standards to comply with the Perkins Act through using measures that do not require information from testing or assessment, or by deferring adoption to their local programs.
   - These states enroll about one-eighth of all high school students.

Types of Skills Assessed

State assessment policies for vocational education cover the four types of skills described earlier—academic, vocational, generic workplace, and broad technical skills (see box 1-A).

Vocational skills—in other words, job-specific skills—are the type most commonly assessed, followed by academic skills. All 43 states with an occupational assessment policy (those in the first three categories above) have a policy for requiring or encouraging assessment of vocational skills.

In the first three categories, 31 states also have in place a policy for assessing the academic skills of vocational students. In addition, all eight states in the fourth category have policies for assessing the academic skills of vocational students. States have apparently responded rapidly to the Perkins

Act requirements for academic outcome measures. The OTA survey was conducted in 1992-93, the school year when most states began implementing standards and measures as directed under the Perkins Act.

Far less frequently addressed are generic workplace skills. Only seven states in the first three categories had policies for assessing generic workplace competencies, and these assessments were typically administered along with assessments of vocational skills rather than being conducted separately.

Broad technical skills are scarcely addressed at all in state policies; only one state emphasized assessment of these skills in the 1992-93 school year.

Use of Assessment Results

States use the results from assessments of occupational skills differently than they use results from assessments of academic skills, according to respondents to the OTA survey. For occupational skills, testing and assessment information are used most often to evaluate instruction and progress in learning, or assess student attainment for course or program completion or certification. The second most frequent use is for accountability, including accountability under the Perkins Act. The third most frequent use is for making decisions about the improvement of courses, programs, or schools.

For academic skills, information from testing and assessment is used most often to meet accountability requirements included under the Perkins Act; second most often for student assessment or credentialing; and third most often to improve programs, courses, or schools. Virtually no assessment components for either academic or occupational skills are being used to meet accountability requirements other than those under the Perkins Act.

There are two important conclusions here. First, for both academic and occupational skills, **the least likely use of testing and assessment information is to improve programs—even though this is the main purpose of the Perkins performance standards.** The reason is unclear; perhaps the information being used is not in a useful form for program improvement or there may be a lack of knowledge about how the information can be used and experience in doing so. Also, data from performance standards will not start to become available in most states until 1993-94, after the first year of operation.

Second, **information about students' academic skills is substantially more likely to be used for purposes of accountability alone than it is for assessing student progress or improving programs.** This finding suggests that state policies for assessing academic skills may have been adopted primarily to comply with the Perkins Act. This indicates that either policies of testing and assessment for academic skills are still in the early stage of implementation, or the academic information being obtained is even less useful than the occupational results for improving programs.

Strategies for Obtaining Assessment Information

Strategies to obtain assessment information vary among states and according to whether academic or occupational skills are being tested. In general, assessments of occupational skills are much more closely tied to local vocational curricula and instruction than assessments for academic skills.

To obtain information about academic skills of vocational students, most states have chosen to use scores from their centrally administered statewide testing program rather than develop new, locally based strategies more closely related to their vocational programs. Of the 31 states that test for academic skills, most use either a state minimum competency exit examination or another test administered statewide at a particular grade level; 26 of the 40 different state academic testing components reported in the OTA survey

were of these two types. Typically these exams are administered in the 9th, 10th, or 11th grade; however, some states are using test score information from the 8th grade. Under these circumstances, it is very hard to see how one could use the information about students' academic skills to improve vocational programs. Most vocational courses are in the 11th and 12th grades.

In contrast, information about occupational skills comes from assessment programs that are either tied to students' completion of a course or course sequence, or are ongoing in the vocational curriculum; 31 of the 54 occupational testing components reported in the OTA survey fit this model.

The relationship between the occupational assessment program and the vocational curriculum varies significantly by state. States that encourage occupational testing—those in category 1 above—strongly tend to focus their testing on completion of courses or course sequences (mostly the latter). States that mandate occupational assessment—category 2—are split between those that focus on the course completion and those in which assessment is ongoing. States that encourage occupational assessment—category 3—tend to use ongoing assessment.

Questions arise as to how states will coordinate dissimilar information from academic and occupational testing to carry out the new program improvement requirements.

EMERGING STATE RESPONSES TO PERKINS PERFORMANCE STANDARDS

The OTA survey also asked states about their plans for expanding, contracting, or continuing their current assessment policies for all four kinds of skills over the next 3 years. Questions elicited details on new components to be added, changes in skills to be assessed, and populations and programs to be tested or measured. The states were also asked about the extent to which they are responding in their expansion plans mainly to the Perkins mandate for performance standards, state educational reform initiatives, state workforce initiatives, or other sources. Questions were structured to determine whether the plans are definite or tentative.

Expansion of Testing and Assessment

The results show that states clearly are planning substantial expansion of their assessment programs and policies. Most of the increase will be due to expansion of existing testing and assessment components, although some states will also add new components. Forty-eight of the 92 testing and assessment components currently in effect are slated for expansion by 1995 and 20 new components will be added. The remaining components will stay the same or nearly the same.

The nature of the changes proposed varies greatly. Some states are planning only minor additions or modifications to existing programs, while others are planning comprehensive new systems. Ohio, for example, will implement an ambitious expansion of its 28-year-old Ohio Vocational Competency Assessment Program (OVCAP). OVCAP will be expanded to include three new tests for generic workplace skills from the Work Keys system (which the state piloted in 1992-93), along with new or revised competency-based tests in at least 63 occupational areas. Ohio is also changing its policy to strongly encourage competency-based tests in all local programs; in the past, these tests have been made available only to local programs that requested them. In addition, the state office of vocational education will obtain scores of vocational students on the Statewide Ninth Grade Proficiency Exam. At the other end of the spectrum, several states are planning only to increase the number of occupational areas covered by their current assessment programs.

Changes in Written Testing and Performance Assessment

Several states intend to shift to written testing for occupational skills and away from

assessment. This trend may be a significant and apparently unintended consequence of the Perkins Act mandates for performance standards.

There will be eight new components of written testing for occupational skills among the states and three new components of assessment; however, two existing occupational assessment components will be eliminated for a net gain of one new component of assessment and eight of written testing. Essentially this means that occupational skills will be measured more like academic skills are now, largely through short-answer tests.

A major issue is how this shift toward written testing will affect the character and content of vocational education. Vocational programs are one of the last places where written testing would seem to provide better information than performance assessment for most purposes. For example, the results of performance assessment consistently have been found to be more directly related to competence on the job than competency testing.

Research also has shown that the imposition of high-stakes, standardized testing for basic skills can narrow the range of skills taught by teachers and learned by students in regular academic classrooms. **Presumably the same kinds of effects could occur in vocational education over the long term, with the content of teaching and learning shifting away from development of more complex skills toward acquisition of factual knowledge, and away from active student production toward classroom lecture. The effects on the character of vocational education could be profound.**

There is very little good research on the instructional effects of written testing versus performance assessment in learning environments like those found in vocational programs, so it is not possible to come to any firm conclusions about the issue. One of the best known pieces of research is an article on training in the military, which gives examples of dramatic effects on learning goals and activities that occurred when performance assessment was substituted for written testing in training environments. When this happened, classroom lecturing on topics such as the muzzle velocity of guns greatly diminished, desks were removed, and students spent much more of their time repairing equipment that they were supposed to be learning how to repair. The performance of students in repairing equipment sharply increased.

It is ironic that written testing methods appear to be expanding in vocational education at the very time questions are being raised about the effectiveness of standardized testing in the rest of education, and experimentation with performance assessment is flourishing. One of the reasons for this is to assess more complex cognitive skills than can be assessed with written testing. Another important issue is how states will coordinate dissimilar information from academic and occupational testing for purposes of program improvement.

Changes in Skills Assessed

State policies are also changing with regard to the types of skills to be measured (see box 1-A). **The greatest expansion will occur in testing or assessment of vocational skills—the job-specific skills conventionally taught in vocational programs.** Testing for vocational skills will expand in 35 of the 50 state components that addressed these skills in the 1992-93 school year. In addition, five new components for vocational skills will be added by 1995.

Assessments of generic workplace skills are expected to increase at a high rate from 8 components to 22, although 17 of the 22 will assess generic workplace skills in combination with other occupational skills.

Assessment of broad technical skills will expand five-fold but will still be the smallest category of testing. There are several reasons for this. One is the lack of clarity about the nature of these skills and the scarcity of instruments and

methods for assessing them. Another is that more emphasis has been placed on integrating academic and vocational education than on reorganizing vocational curricula around broad technical skills. Interest in broad technical skills is mounting, however. The 1990 amendments identified "applied technology," or applied academics, as a theme for vocational education and stressed the need for vocational students to learn about "all aspects of industry." The strongest indication is in the legislation pending before Congress to develop voluntary national industry skill standards, where emphasis would be placed on orienting skill training around clusters of broad occupational skills; supporters contend that this will improve the quality of training, with long-term payoffs for workers and the economy as a whole.

The area with the slowest growth of testing and assessment will be academic skills. Ten states will add academic skills for the first time, but in states that already assess academic skills, expansion is planned for only 17 of 41 components; the other 24 components are not slated for change. This is somewhat surprising, in light of the new Perkins provisions, especially the directive for performance standards to measure gains in student academic skills over time. Test score gains are generally much more difficult to measure than achievement at a single point in time, especially in a mass testing program. It is hard to see how, as many states are doing, test score data from a minimum competency exam given at one point in time or from a state test administered in a single grade can be used to show gains in student academic skills.

Effects of New Policies on Vocational Reform

A key issue is how these state changes in testing and assessment changes will promote or impede the other reform goals of the revised Perkins Act, such as integrating academic and vocational education and broadening preparation in technical skills. **It may well be the case that academic and vocational education are being driven further apart rather than closer together by the responses of states to performance standards.** Resolving this problem will require efforts by the states to develop assessment methods that are more compatible and consistent with the goal of academic integration.

As noted above, states are using different approaches to measure academic and occupational skills. Most states are relying heavily on norm-referenced, standardized tests to measure academic skills, and on locally adapted and criterion-referenced tests to assess occupational skills. Testing for academic skills is also predominantly written, while occupational skills are being measured through a mix of performance assessment and written testing. Academic testing is centralized and conducted statewide at fixed grade levels, while occupational testing is highly decentralized and tied to the local structure of course and program completions. In addition, most of the academic testing occurs in the 9th, 10th, and 11th grades, while the majority of occupational testing occurs as students complete courses or programs, mostly in the 11th and 12th grades. It is hard to see how academic test information collected in grades 9 through 11 can be used to monitor and/or support the integration of academic and vocational education at grades 11 and 12.

In addition, the information from standardized academic tests is often in the form of "rank in class," while the information on occupational skills may be much more performance oriented and competency specific. It is hard to see how such information can be used systematically for the highly localized school-by-school, program-by-program, and teacher-by-teacher nature of the efforts required to integrate academic and vocational education.

In short, there are formidable problems in reconciling and properly interpreting test scores for purposes of monitoring and improving vocational-academic integration. Inattention to these difficulties could lead to serious misuse of test

results. If the content of the statewide academic tests does not reflect the academic content of the programs being integrated, the tests results could give a misleading impression of the success of the efforts.

DEFINING AND ASSESSING BROAD TECHNICAL SKILLS

The implications of organizing vocational programs around broad technical skills are far reaching. Large segments of vocational education are now competency based and oriented to preparing students for specific jobs or job areas. Broad technical skills, by contrast, tend to be clustered around an industry or industry group—examples include health care, financial services, manufacturing, hospitality, and agribusiness. Broad technical skills could even be defined to include historical knowledge of the social and economic development of the industry. Thus, organizing vocational education around broad technical skills could direct more of the vocational education effort to preparing people for careers and on-the-job learning, rather than for specific entry-level jobs.

Data from the OTA survey indicate that organizing vocational education around broad technical skills is a relatively low priority among state and local programs; most vocational programs continue to be oriented toward occupationally specific competencies. Only 1 state assessment component out of 92 current components is oriented primarily to broad technical skills; by comparison, 51 components focus on vocational skills. Only four more components in broad technical skills are planned for the next few years.

A major reason for this low priority is the lack of existing models to illustrate what broad technical skills are and how they can be taught. There are no clear alternatives to the job competency model, which drives prevailing perceptions of the nature of skill, provides the basis for developing curricula and tests, and generally frames vocational education and much of the training enterprise. Concepts of broad technical skills must be defined before assessment programs and methodology can be developed and validated.

As a first step toward developing concepts of broad technical skills, OTA has identified five alternative approaches, founded on substantially different assumptions about the relationships between general and specific skills and between foundation and more advanced skills. These five alternatives are vocational aptitudes, core occupational skills, occupational maps, design and technology, and cognitive skills.

■ Vocational Aptitudes

Vocational aptitude methods reflect a theory that people perform best in jobs for which they have strong abilities, and that these abilities are identifiable through tests. Vocational aptitude tests have been developed by commercial publishers, the military, and others; these tests are used to select people for jobs or training programs or to guide career counseling. They are developed by postulating general and specific abilities that might be good predictors of performance or career outcomes in a range of occupations, and then selecting tests measuring those abilities. Certain domain-specific abilities used in the final test can be viewed as definitions of broad technical skills.

A good example of this type of test is the Armed Services Vocational Aptitude Battery (ASVAB). The ASVAB consists of seven major content areas of general academic and intellectual ability and three content areas of technical ability. The general abilities include verbal comprehension, arithmetic reasoning, and coding speed. The technical abilities are mechanical comprehension, electronics information, and auto and shop information. The three measures of technical skills have been shown through validation research to be significantly related to performance.

Occupational Maps

A second concept of broad technical skills is emerging in some of the industry skill standards projects being supported by the U.S. Departments of Education and Labor. Grantees are urged to organize their standard-setting efforts around "broad occupational areas." The skill definitions are task-oriented, as in the job competency approach used by V-TECS, but are defined across a broad range of similar jobs within an industry.

The American Electronics Association, for example, is considering a structure for its skill standards project that would organize 60 percent of electronic industry jobs that do not require a bachelor's degree into three areas: manufacturing specialist, pre- and post-sales analyst, and administrative/information services specialist. The mapping process involves identifying the major purpose of work in each broad area and defining a limited number of critical functions. For each of these critical functions, generic activities and explicit criteria for good performance are specified. These activities and criteria provide a clearer basis for setting performance-level standards than the job competency method. Categories of knowledge and cognitive skills underlying job performance can be defined as well. For example, for manufacturing specialists, the purpose of the job is to develop, manufacture, deliver, and improve electronics products and processes that meet or exceed customer needs. The initial functions are to:

- ensure the production process meets business requirements,
- initiate and sustain communications,
- establish customer needs,
- determine design manufacturability,
- use human resources to manage work flow,
- select and optimize equipment to meet requirements, and
- make products that meet customer specifications.

Core Occupational Skills

In vocational education, the basic approach to broadening technical skills has been to group vocational programs into clusters of occupations and adopt a guiding common core of required occupational knowledge and skills. At the introductory levels of instruction, all students in all occupational areas take the same courses, which are directly organized around the core skills. At each more advanced level, students take courses in one of the cluster areas. Instruction is organized around the same set of core skills but they become more specialized and embedded in more specific occupational content. In New York, where core occupational skills strategy has been implemented, the skill areas are: a) personal development, b) social/economic systems, c) information skills, d) resource management, and e) technology.

Design and Technology

A fourth approach focuses on development capability for designing technological systems. Advocates assert that all students would benefit from becoming proficient in the design of systems, not only those headed for work or 2-year college programs. The view is that the development of proficiency in designing and building technological systems and learning about technology as a product of human civilization should begin in the early grades. A few such content-oriented concepts could provide a competency-oriented but comprehensive definition of broad technical skills.

Some design and technology courses are taught in the United States at the high school and college levels, but at the high school level the concept has been developed much further in Great Britain and other foreign countries. Over the past 20 years, the British have developed a sequence of courses that now are among the 10 major strands in the new national curriculum, along with mathemat-

ics, science, foreign languages, and other subjects. Design and technology is taken by boys and girls and high and low ability students, beginning in 1st grade. It is conducted to appeal both to students who are interested mainly in the humanities and those who are more scientifically or technically inclined.

Design is the more significant of the two components; students become designers. They acquire the procedural knowledge of learning how to design; the capability to communicate complex ideas with clarity, confidence, and skill; and the conceptual knowledge and understanding of materials, energy, aesthetics, and people and their needs required to create and build effective technological systems.

The procedural knowledge of learning how to design involves learning how to think concretely in complex situations, make choices, and use knowledge to produce better designs. Students weigh the desirability of alternative designs from social, economic, aesthetic, and human standpoints, as well as from the perspective of producibility.

Cognitive Skills

A fifth approach defines broad technical skills in terms of cognitive skills. This approach is based on research from cognitive science that identifies skills needed to troubleshoot equipment and solve problems in a range of occupations, apprenticeship situations, and academic learning.

Much of this research focuses on explaining the differences between the cognitive skills of experts and novices. Research has shown that people who are expert in a domain have acquired large collections of schematically and meaningfully organized factual, conceptual, and procedural knowledge that bears on the technological devices or complex systems in their field. Much of this knowledge is highly specific to these devices and the contexts in which they are used. These structures of knowledge enable experts to understand the complex relationships necessary for skilled performance. Experts differ profoundly from novices in the speed and flexibility with which they can a) access these structures of knowledge, and b) think "metacognitively"—that is, set goals, apply procedural skills flexibly, and learn from experience in solving problems.

Cognitive skills are acquired in stages. Initially a skill is heavily dependent on declarative, or verbal, knowledge. In the declarative stage, the learner either encounters or is taught the facts and procedures relevant to executing a particular skill. These facts are stored in memory as "statements," which can be verbalized and recalled one-by-one in the steps required by the cognitive skill. The second stage is skill automation, or compilation of the cognitive skill. In this process, the factual and conceptual knowledge acquired in the declarative stage is transformed gradually into a highly organized procedural form that can be accessed and used with minimal conscious reasoning activity. The third stage is skill refinement, or proceduralization. In this stage, performance of the skill is speeded up by weeding out nonessential steps and strengthening associations between possible occurrences and effective responses. As a result, performance of the skill becomes much more sensitive to small but critical situational differences, and the flexibility of response to unexpected situations or new data greatly increases. This model has been applied in a wide range of domains, from power plant control to financial planning to tennis.

These basic concepts from cognitive science begin to suggest how the various approaches to broad technical skills are related to each other; they also begin to point the way toward potentially more powerful definitions of broad technical skills. The concepts imply that broad technical skills could be defined in terms of the attributes of thinking and performance that enable individuals to perform in expert-like ways within suitably defined occupational domains in comparison to individuals who are not so expert.

Broad technical skills might be described as skills that are deep in selected areas, but robust

and flexible across broader domains. They consist of tightly integrated structures of contextual, conceptual, and procedural knowledge that are demonstrated and expressed through a variety of verbal, visual, behavioral, and other more tacit ways.

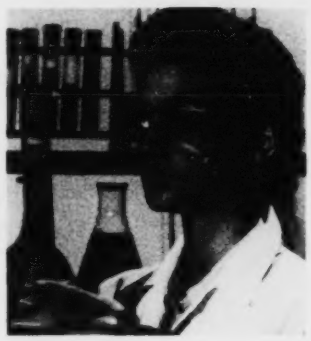
The ability to utilize expertise in novel situations and in new domains needs to be measured by both capacity for responding to new tasks and evidence of general learning aptitudes.

The Evolution of Performance-Based Accountability Requirements in Federal Vocational Education Law

2

In 1990 Congress enacted amendments to the Carl D. Perkins Vocational Education Act that required states to develop and implement standards and measures for the outcomes expected of vocational education students. These standards and measures are to be used to evaluate the effectiveness of local vocational education programs and to stimulate "program improvement" efforts where deficiencies are found.

These provisions for performance standards, which states and local institutions are now implementing, are the most recent stage in the evolution of accountability requirements in federal vocational education law. Accountability means different things to different people, but in the federal context it usually entails some or all of the following aims: assuring that federal funds are used for their intended purposes and beneficiaries, determining whether the federal investment is yielding results, generating political support for federal aid, encouraging more effective programs, and acquiring information for planning and policy decisions.[1] In the past, various tools have been used to achieve accountability, including content and staffing standards; state plans and assurances; detailed guidelines for program administration and service delivery; mandates for local, state, and federal evaluations; federal and state monitoring; and assessments by outside bodies.



[1] Another definition of accountability is that: "It is a process to help people who expect specific benefits from a particular activity (and whose support is essential to its continuation) judge the degree to which the activity is working in their interest so that they might sustain it, modify it, or eliminate it." See Paul T. Hill et al., "Pandora's Box: Accountability and Performance Standards in Vocational Education," paper prepared for the National Center for Research in Vocational Education, December 1992, p. 9.

This chapter seeks to illuminate congressional intent in the 1990 amendments by analyzing the evolution of accountability and evaluation requirements in federal vocational education law and related statutes. The first section of the chapter summarizes the current accountability requirements of the Perkins Act. The second section traces the legislative history of accountability through several decades of vocational education laws. A final section identifies changing themes.

THE CURRENT REQUIREMENTS

The Carl D. Perkins Vocational and Applied Technology Education Act Amendments of 1990, Public Law 101-392, went well beyond a "... routine, business as usual reauthorization ..."[2] by substantially revising the Perkins Act of 1984, the main law governing federal aid to vocational education. Among the more consequential changes were interlocking revisions in the law that have tipped the balance of accountability provisions from funding and planning requirements to the direct assessment of students' outcomes.

Rationale for Outcome-Based Accountability

Several developments spurred the 101st Congress to move toward outcome-based accountability in the Perkins Act. First, national studies indicated "... the need for a massive upgrading in the quality of secondary vocational education." The school reform movement of the 1980s had largely "... bypassed vocational education."[3] Concern was prevalent among business leaders that many high school graduates (including, but not limited to, vocational program graduates) lacked the academic, employability, and occupational skills needed to compete in a global economy. And all too often, the poorest quality vocational programs were found in schools serving low-income students.[4]

Second, evidence suggested that the 1984 Perkins Act had not been as effective as hoped.[5] The congressionally mandated National Assessment of Vocational Education concluded that the Perkins Act was a weak mechanism for improving program quality and increasing access of special populations to good programs; the study suggested performance indicators as a possible way to strengthen the act.[6]

Third, attitudes about how to judge the effectiveness of human resource programs had changed considerably since 1984. "Outcome-based" systems of accountability, which placed relatively less emphasis on whether programs were adhering to procedural requirements and relatively more on whether they were producing the intended results, were gaining popularity at the federal, state, and local levels.[7]

Federal precedent for outcome-based evaluation already had been established through 8 years of experience with Job Training Partnership Act (JTPA) performance standards, new requirements for performance-based evaluation in the Job Opportunities and Basic Skills (JOBS) training program for welfare recipients, and new "pro-

---

[2] Augustus F. Hawkins, *Congressional Record*, U.S. Congress, House, 101st Cong., 1st sess., May 9, 1989, H1709.

[3] John G. Wirt, National Assessment of Vocational Education, testimony at hearings before the House Committee on Education and Labor, Subcommittee on Elementary, Secondary, and Vocational Education, Mar. 7, 1989, p. 5.

[4] Ibid., p. 7.

[5] Among the studies cited as influential by House and Senate Committee reports were U.S. Department of Education, National Assessment of Vocational Education, *Final Report, Volume 1: Summary of Findings and Recommendations* (Washington, DC: 1989); U.S. General Accounting Office, *Vocational Education: Opportunity to Prepare for the Future*, GAO/HRD 89-55 (Washington, DC: 1989); and American Society for Training and Development, *Training America: Learning to Work for the 21st Century* (Washington, DC: 1988).

[6] National Assessment of Vocational Education, op. cit., footnote 5, p. ix.

[7] Richard N. Apling, Congressional Research Service, Education and Public Welfare Division, "Vocational Education Performance Standards," 89-440 EPW, July 6, 1989, pp. 3-4.

gram improvement'' requirements in the Chapter 1 legislation for disadvantaged children. Vocational education seemed a good candidate for this approach, in light of its extensive use of competency-based curricula and its ''somewhat ill-defined tradition'' of examining such outcomes as job placement and employer satisfaction.[8]

Fourth, as the Perkins reauthorization was being considered, the National Governors' Association and the administration were engaged in a broader discussion about the desirability of national standards for what students should know and be able to do in key academic subjects.[9] Around the same time, business panels were beginning to consider the merits of national industry-based skill standards.[10]

Amid this climate, Congress reached a conclusion about vocational education: ''What was acceptable as good or adequate before is not acceptable today if our country is to compete economically.''[11] Not only did vocational education need to be improved on its own terms, wrote House members, but it also needed to be ''. . . tied in much more closely with academic course work . . .'' and with the reforms occurring across the public educational system.[12]

The resulting legislation had several interrelated aims: to target ''. . . money more carefully on programs that produce results . . .,'' to integrate academic and vocational education, to improve programs for special populations of students, and to ease regulatory burdens.[13] Congress seemed to view the law's major provisions as an interdependent system of checks and balances that together would ''. . . reassure critics that the legislative supporters of vocational education were serious about quality.''[14]

Accountability based on standards and measures appears to have been an important component of this system of checks and balances—a tradeoff for relaxing or eliminating certain procedural requirements and funding set-aside. ''Instead of spending money and staff time regulating local programs,'' one senior congressional staff person explained, ''state agencies will now be free to concentrate on securing results.''[15]

Overview of 1990 Accountability Provisions[16]

STATE STANDARDS AND MEASURES

At the heart of the new accountability provisions is a requirement for state boards of vocational education to develop and implement, by September 25, 1992, ''. . . a statewide system of core standards and measures of performance for secondary and postsecondary vocational education programs . .'' (section 115).[17] Congress

[8] E. Gareth Hoachlander, ''Performance Based Policies in Vocational Education,'' paper prepared for the Research Conference of the Association for Public Policy and Management, Seattle, WA, October 1988.

[9] This debate followed the 1989 adoption of six National Education Goals as the Governors and President Bush sought ways to measure progress toward the Goals.

[10] Commission on the Skills of the American Workforce, *America's Choice: High Skills or Low Wages* (Rochester, NY: National Center on Education and the Economy, June 1990).

[11] U.S. Congress, House Committee on Education and Labor, ''Applied Technology Education Amendments of 1989,'' H. Rept. 101-41, Apr. 28, 1989, p. 4.

[12] Augustus F. Hawkins, *Congressional Record*, U.S. Congress, House, 101st Cong., 1st sess., Jan. 3, 1989, E22.

[13] John F. Jennings, ''Congressional Intent,'' *Vocational Education Journal*, February 1991, p. 18.

[14] Hill et. al, op. cit., footnote 1, p. v.

[15] Jennings, op. cit., footnote 13, p. 19.

[16] Section references in this part refer to sections of the Carl D. Perkins Vocational and Applied Technology Education Act (20 U.S.C. 2301-2471) as amended by Public Law 101-392.

[17] For the 1990 amendments, a measure means ''a description of an outcome,'' while a standard means ''the level or rate of that outcome.'' See 32 CFR 400.4.

intended that these systems would "... apply to all the programs assisted under this Act ..." and be developed with the advice of a broadly representative state committee of practitioners.

The state systems are required to include a minimum of two outcome measures:

1. a measure of learning and competency gains, including student progress in basic and more advanced academic skills; and
2. measures for one or more of the following—job or work skill attainment or enhancement; school retention or secondary school completion; or placement into a job, additional training or education, or military service.

By specifying outcomes that encompass academic improvement, dropout prevention, and higher education enrollment, as well as job preparation, Congress endorsed a broad view of the purposes of vocational education. The state systems are also required to include incentives or adjustments that encourage service to special populations (and are consistent with individualized education plans for disabled students).

LOCAL EVALUATION AND PROGRAM IMPROVEMENT

The Perkins standards and measures are designed to derive their real accountability "teeth" from their interaction with new requirements for local evaluation and program improvement (section 117). Beginning with school year 1992-93 and each year thereafter, every local recipient of Perkins Act basic grant funding must use the statewide standards and measures to evaluate annually the effectiveness of its vocational programs. This annual evaluation must also assess how well programs are providing students with strong experience in and understanding of the industry they are preparing to enter. Local recipients may use Perkins Act funding, to a reasonable and necessary extent, to conduct the mandated evaluations.[18]

ACCOUNTABILITY FOR SPECIAL POPULATIONS

The 1990 amendments also changed the approach to accountability for special populations. As summarized by one congressional sponsor:

> For 30 years we have talked about access [of special populations] and rightfully so. . . . When we began writing this bill, however, we asked the question: Access to what? And if we could not answer "access to quality" or "access to excellence," then access was not good enough.[19]

Toward this end, most of the funding set-asides—the mechanism favored since 1968 for serving disadvantaged and disabled students and others with special needs—were eliminated, as were some of the more restrictive process requirements governing services to special populations. In exchange, the 1990 law revised the basic grant funding formula to concentrate more dollars in poor schools and to target federal support on vocational programs that were coherent and sequential, integrated with academic education, and sizable enough to promise meaningful results—in other words, programs with features likely to be effective for *all* students.

STATE PLANNING AND ASSESSMENT

Two other state-level accountability mechanisms deserve mention. First, the 1990 amendments continued the longstanding requirement for states to develop 5-year plans, subject to federal approval, that describe and make assurances about their vocational programs (section 113). Second, prior to submitting their plans, states must conduct an assessment of program quality "... using measurable objective criteria developed by the state board ..." (section 116). This assessment is intended to provide baseline infor-

[18] 32 CFR 403.191.

[19] William F. Goodling, *Congressional Record*, U.S. Congress, House, 101st Congress, 1st sess., May 9, 1989, H1703.

mation about student needs and existing program capabilities for use in state planning.

Congress considered but rejected provisions that would have encouraged or required expansion of the performance standards and measures adopted by the states into a national system of performance standards and reporting for vocational education. The issue was resolved by including in the legislation a study of the comparability across states of the standards adopted and the quality of information available about outcomes. In effect, the study suggests that in the future Congress might consider expanding the performance standards adopted by the states into a national system.

NATIONAL ACTIVITIES

The mandates for outcome-based evaluation at the state and local level are reinforced by complementary activities at the federal level. The law directed the Secretary of Education to:

- provide technical assistance to the states as they develop their systems of standards and measures (section 115);
- submit a report to Congress describing and analyzing state systems and evaluating the degree of comparability across states (section 115);
- consider the implementation of program evaluations and improvements when approving state plans (section 114); and
- conduct research on the development and implementation of performance standards and their effects on student participation and student outcomes, especially for special populations (section 402).

The amendments also authorized a new national program with future implications for state standards and measures: the Business and Education Standards program (section 416). Under this authority, the Secretary of Education has made grants to trade associations, labor organizations, and comparable national organizations to develop national standards for competencies in various industries and trades.[20]

SIGNIFICANCE OF THE NEW PROVISIONS

Realizing, perhaps, that the requirements for standards and measures were breaking new ground and that the cooperation of states and local recipients was vital to the success of this endeavor, Congress took a cautious and incremental approach, remarkable as much for what it does *not* require as for what it does.

First, Congress chose *not* to link vocational education performance standards to funding or other incentives and sanctions, as is done in the JTPA. No authority exists for states "... to apply sanctions in connection with the utilization of measures and standards at the local level."[21] Furthermore, the mild consequence that was attached to evaluations—state intervention through joint program improvement plans—was not intended "... to be punitive in nature, but rather to encourage an infusion of resources from the state for programs that are in need of assistance and improvement."[22]

Second, the legislation did *not* authorize the Secretary to issue national standards and measures, but rather gave states considerable flexibility to develop their own systems. Only two standards are actually required by law, one for student learning and the other for attainment of specific outcomes. Within these broad parameters, states could choose the particular measures for each outcome, add more outcomes, and develop different standards for secondary and postsecondary programs.

---

[20] This program is being operated in tandem with a similar Skill Standards and Certification program in the Department of Labor.

[21] House Committee on Education and Labor, op. cit., footnote 11, p. 14.

[22] U.S. Congress, Senate Committee on Labor and Human Resources, "Carl D. Perkins Vocational Education Act Amendments of 1989," S. Rept. 101-221, Nov. 21, 1989, pp. 22-23.

Third, it was not the intent of Members of Congress "... to set themselves up as the remote, ultimate judges of local processes."[23] Thus, the local level, *not* the federal level, is the primary arena for conducting evaluations, reviewing evaluation data, and carrying out program improvement (with the involvement of the state, if needed). "The Act does not contemplate requiring a recipient to submit the results of the evaluation to the Secretary,"[24] nor are the evaluation requirements structured to produce a national database on program effectiveness. National information needs are to be met through other mandated studies and research.

Fourth, recognizing that students enroll in vocational programs for different reasons, the state system of standards and measures "... is not necessarily intended to apply in its entirety to each individual within the program."[25]

LEGISLATIVE HISTORY

Every major vocational education law since 1917 has included accountability requirements. The form, scope, targets, overseers, and level of detail of these mandates have changed considerably over time, however (see appendix A).

The Smith-Hughes Act

Although the Smith-Hughes Act of 1917[26] did not require program evaluations, it introduced three principles that laid the ground work for accountability and evaluation requirements in later laws. First, Smith-Hughes instituted the concept of using federal funds as a carrot and a stick to stimulate improvement in vocational education and influence state and local policy. Given the newness of the field, improvement was defined mostly in terms of growth: more professional teachers, more and better equipment and facilities, and longer instructional time.

Toward this end, states that desired Smith-Hughes funding had to establish qualifications for vocational teachers, supervisors, and directors paid from federal funds, and minimum requirements for plant and equipment in federally supported vocational classes. The law also prescribed minimum instructional hours per week and per year and required that half the time in trade, industrial, and agricultural programs be spent in practical work experience.[27]

These requirements gave federal administrators a tool that they could wield aggressively to shape the structure and content of vocational education.[28] In fact, federal monitoring of state compliance with federal directives was a primary accountability mechanism during the program's early years.

Second, Smith-Hughes introduced the state plan for vocational education, thereby inaugurating a "... clear-cut, systems management or accountability model which became commonplace as a federal strategy in subsequent decades but was novel for that time."[29] Specifically, the act required state boards to develop plans—subject to approval by a Federal Board for Vocational Education (FBVE)—that described

---

[23] Hill et al., op. cit., footnote 1, p. vi.

[24] 57 *Federal Register* 36842 (Aug. 14, 1992).

[25] House Committee on Education and Labor, op. cit., footnote 11, p. 14.

[26] *U.S. Statutes at Large*, vol. 39, part 1, ch. 114, pp. 929-936.

[27] The Smith-Hughes requirements for minimal teacher qualifications continued until the enactment of the 1976 amendments to the Vocational Education Act of 1963. The requirements for equipment, program format, and instructional time were not included in the 1963 act and were formally repealed in the 1968 amendments.

[28] Larry Cuban, "Enduring Resiliency: Enacting and Implementing Federal Vocational Education Legislation," *Work, Youth and Schooling: Historical Perspectives on Vocational Education*, Harvey Kanter and David Tyack (eds.) (Stanford, CA: Stanford University Press, 1979), pp. 80-81.

[29] Ibid., p. 106.

the programs, schools, equipment, and courses to be funded, the instructional methods to be used, the teacher training to be provided, and the qualifications of teachers and supervisors. The effects of this requirement were far-reaching; federal officials produced detailed guidelines for the content of the plans, participated actively in their development and review, and often called for revisions.[30] As one analyst summarized:

> The creation of a State Plan signed by both state and federal authorities, its review by the FBVE and the staff-written regulations that followed made it possible for the Federal Board to influence directly state vocational programs while contributing modest financial support.[31]

Third, the act enunciated one of the key goals of vocational education that would appear (in updated wording) in all subsequent laws and would eventually be used as a standard for measuring program success: "... The controlling purpose of [vocational] education," the legislation read, "shall be to fit for useful employment. . . ."

Vocational Education Act of 1963

The Vocational Education Act of 1963 is generally regarded as the beginning of the modern federal role in vocational education.[32] Enacted amid a period of rising unemployment, Public Law 88-210 attempted to respond to criticisms that vocational education—with its emphasis on the traditional areas of agriculture, home economics, and trades—had failed to keep pace with "... the more sophisticated economy of the 1960's."[33] The 1963 act sought to improve vocational education and provide access to training that was "of high quality" and "realistic" in light of employment opportunities.

The 1963 act introduced the concept of program evaluations. States were required to conduct and give "... due consideration ... [to] periodic evaluations of state and local vocational education programs and services in light of information regarding current and projected manpower needs and job opportunities."[34] Decisions about how to conduct and fund these evaluations and what they would examine were left entirely to the state, so long as some use was made of labor market information.

The law also created an ad hoc national advisory council that would review and recommend improvements in vocational education—the first of several outside bodies charged with studying vocational programs. As with Smith-Hughes, the concept of "improvement" in 1963 was viewed largely in terms of expanded infrastructure and better services.[35] This advisory council was later formalized in the 1968 amendments. It continued in the legislation up until the 1990 amendments, when it was terminated.

Another key provision reserved at least 3 percent of each state's grant for ancillary services, defined to include *program evaluation,* teacher training, demonstration and experimental programs, materials development, and state administration and leadership.

Vocational Education Amendments of 1968

By 1968, federal support had helped fuel tremendous expansion of vocational enrollments

---

[30] Whether this federal influence helped or hindered education is an area of disagreement among the limited body of research about the early years of vocational education. See ibid.

[31] Ibid., p. 105.

[32] Intermediate legislation affecting vocational education included the George-Reed Act of 1929, the George-Deen Act of 1936, and the George-Barden Act of 1946.

[33] U.S. Congress, Senate Committee on Labor and Public Welfare, Vocational Education and National Institute of Education Amendments of 1976, S. Rept. 94-882, May 14, 1976, p. 42.

[34] 77 Stat. 406.

[35] U.S. Congress, House Committee on Education and Labor, "Vocational Education Act of 1963," H. Rept. 88-393, June 18, 1963, p. 6.

and expenditures. Even so, concerns persisted that vocational programs were not adequately preparing students for growth occupations and were maintaining outdated programs.[36] Enunciating a theme that would reemerge in 1990, the House Committee concurred with national advisory council findings that "... any dichotomy between academic and vocational education is outmoded," and that the development of attitudes, basic educational skills, and work habits was as important as training in specific occupational skills.[37]

The Vocational Education Amendments of 1968, Public Law 90-576, sought to address these concerns through revised accountability requirements and other means. The requirement was extended for states to conduct periodic evaluations based on labor market needs and consider the results in program planning. A new authorization of funding for state administrative activities would help support evaluation.

The most significant change in the 1968 amendments was a new provision that limited federal support to programs that could be demonstrated "... to prepare students for employment or ... for successful completion of such a program, or be of significant assistance to individuals enrolled in making an informed and meaningful occupational choice."[38] (Homemaking education was exempted.) The amendments contained no guidance, however, about how local programs might demonstrate these outcomes to the state, nor were they construed by states to mean that a formal outcome-based evaluation was necessary.

Education Amendments of 1976

Approaching reauthorization in 1976, Congress pointed to several dramatic changes resulting from federal support for vocational education, all based on "inputs": increased enrollments, higher expenditures from all sources, construction of area vocational schools, more and better trained teachers, and a greater number and variety of course offerings. As the House Committee noted, however, evidence of program outcomes was sorely missing:

> Measures of the outputs—the success of the program completers and leavers in finding and keeping jobs—are more difficult to find.... [T]here is some scattered evidence of placement success for a number of local programs. But in terms of periodic and extensive reviews of these programs, in terms of the success of their students in obtaining and keeping jobs, little can be found at present. For that reason, the Committee has recommended amendments to provide this type of information for measuring the effectiveness of the programs.[39]

Accountability was found wanting on several other scores. A General Accounting Office report criticized the Department of Health, Education, and Welfare for failing to follow up on evaluation documents submitted by states, for conducting perfunctory reviews of state plans, and for inadequately monitoring state programs.[40] In addition, the lack of reliable national occupational information and local employment data made it hard for states and local recipients to conduct solid evaluations. Perhaps the greatest failing was in follow up of job placement and employer satisfaction,

---

[36] U.S. Congress, House Committee on Education and Labor, "Vocational Education Amendments of 1968," H. Rept. 90-1647, July 8, 1968, p. 2. The criticism of using federal funds to maintain outdated or poor quality programs would surface in subsequent reauthorizations, until the use of funds for program maintenance was restricted somewhat in 1984.

[37] Ibid.

[38] 82 Stat. 1076.

[39] U.S. Congress, House Committee on Education and Labor, "Education Amendments of 1976," H. Rept. 94-1085, May 4, 1976, p. 12.

[40] Ibid., p. 17.

[41] Ibid., p. 20.

which congressional witnesses testified "... was very sporadic and extremely uneven."[41]

Concluding that existing mandates seemed "... to be having little effect,"[42] Congress significantly strengthened the accountability and evaluation requirements in Public Law 94-482, the Education Amendments of 1976. The 1976 law for the first time contained a separate section on evaluation (section 112), with the aim of "... assist[ing] states in operating the best possible programs of vocational education." Within the 5-year period of the state plan, every state had to evaluate the effectiveness of each local program and use the results in revising state programs and plans. For every program purporting to impart entry-level job skills, the evaluations are required to show the extent to which program completers and leavers: a) found employment in occupations related to their training or pursued additional education; and b) were considered by employers to be well trained. These two gauges of effectiveness were specified "... because in [the Committee's] opinion they show most clearly whether persons trained in vocational programs are showing the results of such training."[43] Programs that were "... purely introductory or preparatory to actual job training ..." were excluded, and data was to be collected by sampling wherever possible to reduce the burden.[44]

The 1976 amendments signified a high-water mark for the use of plans, applications, and reports as accountability tools. Public Law 94-482 required a 5-year state plan, developed with the involvement of 10 representative groups; an annual program plan; an annual state accountability report that summarized evaluation findings and described how they were used to improve programs; and local applications. As part of the 5-year state plan, an assessment of current and future state needs for job skills was also mandated.

How well the 1976 evaluation requirements were implemented was a topic addressed by the National Institute of Education (NIE) study of vocational education. "State and local vocational educators objected strenuously to the evaluation requirements ..." and felt overwhelmed by the prospect of implementing the "... new, complex, and costly ..." procedures.[45]

Despite state complaints, the NIE study found that the 1976 evaluation provisions did stimulate improvements in evaluation. In 1976, few, if any, states had adopted evaluation procedures as comprehensive as those called for in the act. Program reviews were the most frequently implemented activity resulting from the amendments. At the same time, the study identified practical problems with the reliability and consistency of the followup data and also found that employer satisfaction data was being collected much less frequently than student placement data.[46] Many of the pieces were in place, however, and the amendments helped draw together and systematize these discrete elements.

Carl D. Perkins Vocational Education Act

The 1984 Carl D. Perkins Vocational Education Act, Public Law 98-524, replaced the Vocational Education Act of 1963. In some respects, it represented the apex of federal prescriptiveness in vocational education, especially regarding services for special populations, which were funded through set-asides totaling 57 percent of state basic grant funding. The remaining 43 percent of

[42] Ibid., p. 38.

[43] Ibid., p. 38.

[44] Ibid., p. 39.

[45] State and local administrators argued that the emphasis on student placement failed to reflect the broad goals of vocational education, overlooked economic forces outside the control of schools, and might diminish service for hard-to-place students. See U.S. Department of Education, National Institute of Education, *The Vocational Education Study: Interim Report* (Washington, DC: 1980), pp. V-5 to V-8.

[46] Ibid., p. V-11.

the basic grant was targeted on "program improvement"—which in the 1984 context meant making programs more responsive to contemporary labor market needs, especially high-technology fields, and updating equipment, curriculum, and staff. The 1984 reauthorization occurred in a climate in which the very existence of high school vocational education was being questioned.[47]

As one of several amendments aimed at improving program quality, the 1984 Perkins Act directed states to develop measures of program effectiveness, such as the occupations to be trained for, the levels of skills to be achieved, and the ". . . basic employment competencies to be used in performance outcomes, which will reflect the hiring needs of employers . . ." (section 113).

The U.S. House of Representatives sought more specific and outcome-based evaluation requirements than those that found their way into final law. The House wanted the states to develop objective standards for the outcomes of occupationally specific programs, which included adjustment factors for local situations, and apply the standards to the approval of local plans and the direction of technical assistance to improve local performance. Although the House Committee took pains to distinguish that these expected outcomes were not performance standards, they foreshadowed the requirements for standards that would be adopted 6 years later.

Foreshadowing the Business and Education Skill Standards program, the 1984 law also required states to establish technical committees, composed of business and labor representatives, to develop inventories of skills for priority occupational fields, which could be used to define model curricula. These technical committees influenced vocational education to adopt the job-competency model, where instruction and testing and assessment are closely tied to the specific skills needed for individual jobs.

Other Influential Statutes

Prior to 1990, Congress had already enacted requirements for performance-based accountability and program improvement in other federal education and training legislation. The approaches varied, but several influenced the Perkins Act amendments. Two of the most important programs are the Job Training Partnership Act and the Chapter 1 Program of the Elementary and Secondary Education Act.

JOB TRAINING PARTNERSHIP ACT OF 1982

The JTPA was a trailblazer in performance-based evaluation and continues to be a strong influence on other federal human resource programs, including vocational education. Unlike vocational education, the JTPA is a wholly federal creation, completely federally funded, and is mostly directed to low-income individuals.

Performance standards are established according to a hierarchical process, starting with the definition of certain broad initial outcomes in the authorizing legislation.[48] For adults in Title II programs, these outcomes are placement in unsubsidized jobs, retention in unsubsidized employment, increased earnings, reduced welfare dependency, and acquisition of employability and basic skills (including receipt of a high school or equivalent diploma). For Title II youth programs, all of the adult outcomes are applicable, plus attainment of employment competencies, secondary and postsecondary school completion, dropout prevention and recovery, and enrollment in other training, postsecondary education, or military service. Based on these broad parameters in the statute, the Secretary of Labor provides further detail, selecting specific outcomes that

[47] U.S. Congress, Senate Committee on Labor and Human Resources, "Carl D. Perkins Vocational Education Act of 1984," S. Rept. 98-507, June 7, 1984, p. 2.

[48] See section 106 of The Job Training Partnership Act of 1982, Public Law 97-300, enacted Oct. 13, 1982; amended Dec. 31, 1982, Public Law 97-404, sec. 1(b).

conform with the statutory intent, establishing core standards and quantitative measures for each one, determining the time period of measurement, and suggesting cost-effective ways for obtaining the remaining data.

In the early years of the program, these standards focused primarily on short-term outcomes attained immediately on completion of training. Recently, the Department of Labor has tried to extend the time period to 13 weeks after program termination.

The law gives governors the flexibility to determine the relative importance of different standards and to prescribe variations in standards based on state economic, geographic, and demographic conditions, as well as characteristics of the populations served. Governors may also prescribe additional standards, develop incentives, and sanction policies, including incentives for hard-to-serve individuals. Local Private Industry Councils make the final translation into bottom-line criteria for service providers.

Unlike vocational education, the JTPA attaches sanctions and incentives to the standards. Programs that fail to meet standards for 2 years, after receiving technical assistance from the governor, are subject to a state-imposed reorganization plan, which may shift funds to another local administrative entity. In certain cases, the Secretary can withhold up to one-fifth of the local recipient's funds for technical assistance.

The JTPA experience with performance standards shows both their benefits and their potential pitfalls. "In conjunction with clearly identified client and service goals, performance standards appeared to have their intended effects of increasing efficiency and accountability," a 1988 study found.[49] At the same time, JTPA performance standards have been criticized for encouraging "creaming"—focusing services on clients who are easiest to place rather than the most disadvantaged—because the prior standards measured success primarily through high placements and low cost.

In summary, while the vocational education standards address some of the same outcomes as the JTPA standards and, like the JTPA, allow for adjustments for conditions, they differ from the JTPA in that they are not tied to funding, other sanctions, or incentives, are not based on national numerical measures, and do not address cost issues.

CHAPTER 1, ELEMENTARY AND SECONDARY EDUCATION ACT

Chapter 1 (formerly Title I) of the Elementary and Secondary Education Act of 1965 was one of the first federal education programs to mandate evaluations of student outcomes. In this and other respects, it has substantially influenced the Perkins Act. Yet Chapter 1 differs from vocational education in many ways, not the least of which is in the evaluation system that has arisen from the 1965 requirement.

Chapter 1 is completely federally funded, supplementary to the regular school program, aimed at a distinct target population, and focused primarily on academic skills. Evaluation in Chapter 1 is governed by a national system, the Title I Evaluation and Reporting System (TIERS). Unlike vocational education, Chapter 1 uses a single outcome measure—the scores of participating students on standardized achievement tests—to evaluate program effectiveness. Local school districts (the recipients of Chapter 1 funds) must conduct pretesting and post-testing of Chapter 1 students and report the scores to the state in accordance with a complex set of TIERS technical guidelines. The state aggregates and reports the test scores to the federal government, which further aggregates the results into a national picture of Chapter 1 student achievement.

Chapter 1 evaluation data took on greater significance with the 1988 enactment of new

[49] Katherine P. Dickinson et al., *Evaluation of the Effects of JTPA Performance Standards on Clients, Services, and Costs* (Washington, DC: National Commission for Employment Policy, 1988), p. 4.

program improvement provisions (Public Law 100-297), which were spurred by concerns about stagnating program quality and were accompanied by a loosening of process requirements in other parts of the legislation.[50] Under the 1988 amendments, local school districts must conduct annual evaluations of Chapter 1 student performance and must implement program improvement plans in project schools that do not exceed state minimum requirements for Chapter 1 aggregate student achievement. If, after a year, school performance still does not improve, then the state becomes involved through a joint improvement plan. The process continues until improved performance is sustained for more than 1 year.

Federal regulations have set a minimum standard for annual test score gains, but states are encouraged to establish higher standards. States and local school districts are also encouraged to evaluate local programs on the basis of "desired outcomes" other than standardized test scores, but there is little incentive to do so, since more outcomes mean additional hurdles for program improvement. Chapter 1 also requires schools to examine the progress of individual children and to conduct a thorough assessment of program needs for children who do not meet minimum standards.

The Perkins Act program improvement provisions drew some key features from Chapter 1: the authority for states to develop standards, the requirement for consultation with a committee of practitioners, and the process for local and state program improvement plans. Unlike Chapter 1, however, vocational education standards do not rely heavily on a single measure and are not tied to a national reporting and evaluation system.

CHANGING THEMES

Several themes emerge from the legislative history of vocational education that help to clarify congressional intent about the new requirements in the Perkins Act for accountability, show the shifts in federal policy on accountability in vocational education, and highlight issues likely to arise during implementation or future reauthorizations.

The mechanisms for accountability in vocational education have changed substantively over time, as the federal government has sought better ways to improve program effectiveness and achieve federal goals in a field with an increasingly strong state and local presence.

- In the early years of the program, from 1917 to roughly 1963, accountability was enforced largely through federal approval of state plans and federal monitoring of state programs.
- State planning peaked as an important accountability tool in 1976 with mandates for multiple plans and accountability reports.
- Since 1963, mandated reviews of the quality of local programs and the access of special populations to them has been a linchpin of accountability. Initially, these evaluations were oriented to the state review of the quality of local programs using criteria other than student outcomes. The initial step toward student outcomes was made in 1976, with requirements for followup information on job placement and employer satisfaction. The 1990 requirement for performance standards sharpens the focus on student outcomes.
- Responsibility for conducting program reviews has shifted in recent years from states to local recipients, bringing the activity closer to those with the greatest stake in the outcomes and the greatest likelihood of using them to revise programs. The 1990 amendments require reviews of the quality of local programs both by the states using their performance standards measures and by local programs themselves. The local reviews are to be conducted annually.

[50] U.S. Congress, House Committee on Education and Labor, "School Improvement Act of 1987," H. Rept. 100-95, May 15, 1987, p. 22.

- Through 1976—or through 1984, on some issues—federal procedural requirements and funding set-asides became more numerous and detailed. In recent years, however, the balance has shifted somewhat away from these mechanisms and toward outcome-oriented standards.
- Throughout the legislative history, the federal government has frequently turned to quasi-independent bodies, such as national, state, and local advisory councils and the National Center for Research on Vocational Education to conduct evaluations, provide technical support, and solicit business community advice on vocational education.

Reasons for Shifts in Accountability Requirements

Several reasons underlie these shifts in the accountability provisions of the federal legislation. The reasons include persistent concerns about the effectiveness and relevance of vocational programs, changing definitions of quality, state and local backlash against federal prescription, and the strong desire of Congress to maintain the viability of the federal program. Still, tying vocational education to student outcomes is difficult because of its multiple goals, which are hard to measure, and variation in priorities accorded them by states and local communities.

One of the most important reasons is that the accountability requirements of the federal legislation often do not seem to Congress to have produced the desired results. Concerns about the effectiveness of vocational education programs and their relevance to labor market needs have persisted through almost every reauthorization of vocational education back at least to the Vocational Act of 1963. As one researcher noted in 1979:

> [After] sixty years of school programs and after billions of federal, state, and local dollars, no legislator, educator, or lobbyist can prove that vocational programs do precisely what federal legislation promised. On the contrary, in the last forty years blue-ribbon committees and government financed studies have pointed out repeatedly serious shortcomings in vocational education allocations, operations, productivity, and impact within schools.[51]

Second, as the field of vocational education has matured, the definition of a quality program has changed from infrastructure (e.g., adequate facilities, equipment, and professional staff), to equity (e.g., access for special populations), to modernization (e.g., market-relevant courses and updated curricula), and eventually to student impacts (e.g., job placement and competency attainment). As definitions have changed, so have the means for enhancing quality. The specific, overall goal of vocational education addressed most often in evaluation requirements in recent years is that of preparing students for employment. From an initial directive in 1963 for evaluations based on "manpower needs and job opportunities," the legislation has become more specific about how progress toward employment goals should be measured. Other goals—such as dropout prevention and academic achievement—have not been targeted for evaluation until quite recently.

Finally, it is telling that policymakers have responded to evidence of shortcomings in vocational education by devising new, different, or stronger accountability mechanisms rather than by eliminating federal support or merging vocational education with job training, as some have proposed. This decision suggests that legislators are committed to maintaining vocational education as a viable system and that they believe the federal government can influence state and local policy even with a very limited share of federal funding.

---

[51] Cuban, op. cit., footnote 28, p. 70.

Balancing Accountability and Flexibility

For much of the 1960s and 1970s, the federal government relied on detailed process requirements to ensure that human resource programs reached the intended beneficiaries and produced the desired results. In the early 1980s, as state and local criticism heightened about burdensome, counterproductive, and overly prescriptive federal mandates, Congress responded by relaxing requirements in many programs, from the Comprehensive Employment Training Act (CETA) to Chapter 1.

Vocational education represented a somewhat special case in this debate. The early years of the program were characterized by a high degree of federal influence and dependence on federal funding. During the 1960s, however, state flexibility actually increased, as the legislation eliminated some rigid funding categories and gave states more discretion over which programs to support and how to use their federal grants. Federal efforts to exert stronger influence in 1968 and 1976 were undercut somewhat, as the federal share declined and states channeled their own resources into programs that were not always consistent with federal priorities. In 1984, as Congress was relaxing requirements in other education programs, it continued to seek stronger mechanisms for enforcing federal priorities in vocational education. The results were mixed.

In 1990, Congress changed course and moved to require outcome-based evaluation in exchange for increased flexibility. The tradeoff, however, was not as tidy as congressional rhetoric suggests. The 1990 amendments, though less prescriptive than the 1984 Perkins Act, are still rather detailed. Although many process and set-aside requirements were eliminated, other new requirements were added, governing funds distribution, program content, and the integration of vocational and academic education. However, the emphasis on performance-based accountability has been cautious compared to the JTPA.

Through several reauthorizations, Congress has also tried to balance tensions that arose between the goals of access and excellence and to ensure that vocational education services to special populations were of high quality. The 1990 amendments went several steps further by eliminating the set-asides and requiring high-quality services for special populations.

As the JTPA experience shows, performance standards do not necessarily resolve the tension between these two goals and may create new challenges. Learning from the JTPA, the 1990 amendments sought to build in safeguards to ensure that standards and measures would include adjustments for serving special populations and that local evaluations would include a review of the progress of special populations.

Cautious Approach

As noted above, the 1990 legislation took a cautious approach to performance outcomes in vocational education. The standards are state-developed, not nationally developed. There is no provision for reporting local evaluations to the federal government, nor are funding sanctions attached to the results. The program improvement process is meant to be helpful, not punitive. There are also several provisions for additional research, technical assistance, demonstration, and data collection regarding implementation of the performance standards.

In short, Congress built in several opportunities to monitor the progress of implementation and keep informed of difficulties that may arise. This suggests that the new accountability provisions are conceived as a first step, to be reviewed carefully before the next reauthorization. This deliberate approach to performance standards is also reflected in other programs, such as the phase-in periods for standards in the JOBS and Food Stamp Education and Training programs.

State Policies on Testing and Assessment

3

As with any new policy, the ultimate effects of the shift in federal policy on accountability to performance standards will depend on how it is implemented in state and local programs of vocational education. The role of the states will be pivotal in the implementation process because of their responsibilities for setting the performance standards and adopting policies of testing and assessment measuring progress.

Beyond the initial definition of the performance standards and measures for a state, successful implementation will require the development of substantial state resources for assessing the academic and occupational skills acquired by students. In many states, new resources will be needed.

As indicated in chapter 2, the legislation requires the adoption of outcome measures for learning and competency gains, and at least one other area of competency attainment or employment outcomes. Most states have adopted sets of standards and measures in at least four or five areas, involving some combination of gains in or attainment of academic skills, gains in and/or attainment of occupationally specific skills, attainment of general work or employability skills, rates of program completion, rates of job placement, and status of employment or further education.[1] In most cases, all except the last three types of outcomes will require information about the performance of individual students from some form of testing or assessment. The National Center for Research in Vocational Education's (NCRVE) recent tabulation of the standards adopted by the states shows

[1] Mikala L. Rahn et al., "State Systems for Accountability in Vocational Education," paper prepared for the U.S. Department of Education, Office of Vocational and Adult Education, December 1992.

that a majority of them at least will require some information from testing and assessment.[2]

Consequently, the eventual success of performance standards in stimulating reform and improvement in vocational education will depend on the resources for testing and assessment within state and local programs. If no resources exist or the quality of the information from testing and assessment is low, conclusions derived will be faulty. The testing or assessment process might not be focused on the most important outcomes of the local programs, or the results may not be dependable because of the instruments employed, how they were presented to students or administered, or how the responses of students were rated or scored. The problems could be random, in which case they might not be threatening to the integrity of the resulting information, or they could be systematic, in which case they would be a problem. Evidence from academic education shows that when the stakes for testing and assessment are sufficiently high, the process of interpreting the results can become highly politicized and the results can be distorted.[3]

This chapter presents the results of a "first cut" effort by the Office of Technology Assessment (OTA) to describe the policies of states on testing and assessment for the academic and occupational skills of vocational education students, and the state plans for expanding their policies on testing and assessment by 1995. The data presented come from a survey conducted by OTA, along with interviews with state personnel and others who are knowledgeable about practices of testing and assessment in vocational education.

The results presented serve three main purposes. One is to describe the range of current practices of testing and assessment in vocational education. Partly, this is done through comparing current practices with practices of testing and assessment in academic education. There are clear differences between the philosophies and origins of testing and assessment in academic and vocational education that are important to keep in mind. The second purpose is to describe the plans of states for expanding their resources for testing and assessment in response to the requirements for performance standards in the Perkins Amendments, and other forces at work at the state and local levels in vocational education. The third is to consider the correspondence, or lack thereof, between the emerging policies of states on testing and assessment and two reform goals: integrating academic and vocational education, and broadening the technical skills around which vocational education is organized.

The testing and assessment policies of states are described in terms of two dimensions. One dimension is the four types of academic and occupational skills, and the other is the extent to which methods of written testing or assessment are emphasized. These categories of academic and occupational skills are important because they are related to the integration of academic and vocational education, and broadening the technical skills around which vocational education is organized. Whether the state policy emphasizes written testing or assessment is important because of the potential effects on the content and character of instruction in vocational education, and because of validity and reliability issues.

For the purposes of this study, *written testing* will be defined to be any method of examining students in which the format of the answers to the questions asked are multiple choice, matching, or some other method of filling in a small blank or selecting the correct response from a given list of responses. Such measuring instruments are *closed ended. Assessment* will be defined as the

---

[2] Ibid., table 2.

[3] U.S. Congress, Office of Technology Assessment, *Testing in American Schools: Asking the Right Questions*, OTA-SET-519 (Washington, DC: U.S. Government Printing Office, February 1992), ch. 2; and Daniel Koretz, "Arriving in Lake Wobegon: Are Standardized Tests Exaggerating Achievement and Distorting Instruction?" *American Educator*, vol. 95, No. 2, 1988, pp. 46-52.

observation and rating or judging of student performances to ascertain academic or occupational skills, following some systematic procedure of measurement.

DESIGN AND METHODOLOGY OF THE STATE SURVEY

The OTA survey was conducted by telephone from February through April of 1993. It had two parts: 1) a series of closed-form questions to obtain basic descriptive information about the individual components of each state's program of testing and assessment, and 2) further discussion with the state respondents to develop a brief, written description of the nature, form, and purposes of each of the components and how they are related to each other. Both the coded questionnaire responses and the qualitative descriptions were sent back to each state for verification of the accuracy and completeness of the information. The respondents were the state directors of vocational education in each of the 50 states and the District of Columbia, or a person designated by the state directors. All 51 responded. Questions were asked only about secondary vocational education.

The survey was organized around the collection of data for each of the major, individual *components* of each state's program of testing and/or assessment, not the program as whole. In a state with a well developed program of testing and assessment, these components could be, for example:

1. a state policy or guideline of requiring local programs to monitor the academic achievement of all students who complete a vocational program through the administration of a commercially available test, such as the Test of Adult Basic Skills (TABE);
2. a state policy or guideline of making a competency-based test item bank available to local programs and strongly encouraging them to construct their own tests for measuring the vocational skill attainments of all students on completion of vocational programs; and
3. a state policy of strongly encouraging local programs to provide all students who complete vocational programs with a profile or cumulative portfolio showing their accomplishments and competencies to use in seeking employment.

Data was collected by components to be as precise as possible about the nature and extent of each state's policies of testing and assessment, and plans for the expansion or contraction of those policies through 1995. These components of testing and assessment are the basic unit of analysis of the study.

Although the data describe only the policies of states on testing and assessment, the information also appears to provide a reasonably accurate picture of policies and practices at the local level in many states. Local conformance is mandated in some states and in many others there are strong traditions of following the lead of the state agency, even if policy is not mandated.[4] In the remaining states, the range of local policies and practices is broad.

The survey was designed by OTA to describe all components of each state's policies and practices of testing and assessment, rather than only the components employed for implementing performance standards and measures. This provides a more accurate basis for describing practices of testing and assessment in vocational education, and examining the effects on these practices of the requirements for performance standards in the federal legislation. Testing and assessment are conducted for many reasons other than performance standards. Insofar as possible, the information collected by OTA was compared

[4] Lawrence Cuban, "Enduring Resiliency: Enacting and Implementing Federal Vocational Education Legislation," *Work, Youth, and Schooling: Historical Perspectives on Vocationalism in American Education* (Stanford, CA: Stanford University Press, 1982), pp. 45-78.

item-by-item with the data on the plans of states for implementing performance standards and measures that have been collected by NCRVE.[5] Followup calls were made to resolve all differences. All data was obtained from the states by telephone but then sent back to them for their review and explicit approval.

For each component, questions were asked in the telephone survey and interviews to determine:

- What skills are assessed within the four broad categories of academic, vocational, generic workplace, and/or broad occupational skills.
- How information resulting from testing and assessment is used by the states.
- Where in the curriculum testing or assessment is conducted (e.g., by grade level, in introductory vocational education courses, or at the end of a sequence of occupationally specific vocational courses).
- Whether all local programs are required as a matter of state board policy or legislation to conduct testing and assessment or are only encouraged by the state to conduct it.
- What resources for testing and assessment are made available to local programs by the state as part of the state's policy on testing and assessment.
- What the state's plans are for expanding, contracting, or adding to their program of testing and assessment through 1995-96.

Practices of testing and assessment in vocational education have not previously been the subject of much research. OTA asked the obvious and most simple questions about testing and assessment practices in vocational education in order to provide basic descriptive information and raise policy questions. The most extensive existing study is the recent NCRVE report on the implementation of performance standards. The NCRVE study lists the performance standards that have been adopted by the states and some of the measuring instruments that will be used.[6] A survey of state practices in competency-based testing has also occasionally been conducted by the Research and Curriculum Unit for Vocational, Technical, and Adult Education of Mississippi State University.[7] The most recent report describes the states that have programs of competency-based testing for vocational skills and make competency-based tests or test item banks available to their local vocational programs. The National Association of State Directors of Vocational Education has also recently completed a one-time survey of the industry skill standards initiatives of states that contains some information on assessment practices.[8]

ORIGINS OF TESTING AND ASSESSMENT POLICY IN VOCATIONAL EDUCATION

The origins and current practices of testing and assessment in secondary vocational education are different from the rest of education in several important respects. The phrase "testing and assessment" that has been repeatedly used above begins to reveal some of these differences. Roughly speaking, there are two related traditions of testing and assessment in vocational education. These two traditions have common origins in the competency-based movement. They differ in the emphasis placed on the need for written testing to provide reliable measurement, as opposed to allowing the use of a broad range of methods of

[5] Rahn et al., op cit., footnote 1, appendix.

[6] Ibid.

[7] National Network for Curriculum Coordination in Vocational Technical Education, Research and Curriculum Unit for Vocational, Technical, and Adult Education, Mississippi State University, "1990-91 State of the Art Report on Statewide Student Competency Testing in Vocational and Technical Education," unpublished report October 1990.

[8] Barbara Border, *Education-Driven Skill Standards Systems in the United States*, prepared for the U.S. Department of Education (Washington, DC: National Vocational Technical Education Foundation, October 1993).

observation and evaluation, some of which may involve a substantial amount of judgment on the part of instructors or others.

The written approach to *competency testing* in vocational education involves the administration of tests in which the questions are keyed directly to specific items of knowledge and skill needed on the job, and the answers are provided by some method of checking off the correct response from a given list of responses. In vocational education, most of these written tests are matching or multiple choice.

The range of *competency-based methods of assessment* used in vocational education is broad. Included is the administration of carefully designed performance exercises, the summary evaluation of projects undertaken by students, organized events or competitions in which students compete for recognition and rewards, and even the subjective rating of regular classroom work within some framework of performance elements, rating scales, and rating procedures. There are also longstanding practices of providing students with profiles and/or encouraging them to accumulate portfolios of their schoolwork to use in seeking employment.

For the purposes of this report, assessment is defined as a process where student responses to a task or variety of tasks over a period of time are carefully observed and evaluated or interpreted according to some set of agreed on criteria or dimensions. The tasks may be presented to the student by the teacher or a test administrator in the form of specific "prompts" or statements of problems to be solved. They could alternatively be initiated by the student in response to general instructions. Students may respond to the prompts or problem situations *on demand*—that is, they are assigned at a certain point in time and must be responded to within a limited period of time; or performance may occur over an extended period of time, as a part of the regular program of instruction and student work.

The differences between competency testing and assessment in vocational education parallel in some respects the current debate in academic education over the future of standardized testing and performance assessment. **There are, however, virtually no traditions in vocational education of reliance on the kinds of norm-referenced, standardized tests of academic skills that are so prevalent in academic education.** Testing in vocational education stems from entirely different origins than standardized testing in academic education. Competency testing in vocational education stems from the competency-based movement in vocational training, while standardized testing in academic education is descended from the mental testing movement that began in psychology around the turn of the century and has resulted in the concept of ability testing.[9]

One of the major differences between the competency-based, written testing done in vocational education and the standardized testing done in academic education is that in vocational education the tests are constructed to measure whether students have the skills needed to perform particular jobs, rather than how their performance compares with other students taking the same test. In the language of test theory, the written tests in vocational education are criterion-referenced rather than norm-referenced.

In properly conducted competency-based vocational education, both the content of the tests and the curriculum of instruction are criterion-referenced—that is, they are derived from analyses of the tasks actually performed by people working in specific occupational areas. The relationship between the job tasks identified and the content of instruction and individual test items is close.

Competency testing and assessment are a key aspect of the concept of open entry/open exit that is followed in many vocational programs. This method of organizing instruction is basic to the

[9] Office of Technology Assessment, op. cit., footnote 3, ch. 5.

competency-based approach, but not all programs are organized in this way. In open entry/open exit instruction, students are tested or assessed when entering programs to determine their skills. They then work to acquire the particular skills required to achieve their employment goals; they leave the program when they have demonstrated mastery of those skills, according to the results of tests or assessments. Students learning alongside each other may be learning very different skills in highly individualized programs of instruction.

The traditions of competency-based assessment in vocational education are older than competency-based testing but stem from the same origins in the competency-based movement. The specific lists of job competencies are employed as content and performance-level standards for observing and recording the capability demonstrated by students in different kinds of assessments. The content standards define the skills that students should demonstrate. There is also usually a categorized scale of performance levels that defines degrees of proficiency in performance. These performance levels are defined according to a scheme showing that the student is, for example, "skilled," "moderately skilled," "unskilled," or "has not been exposed" to a particular task or sets of tasks.

The philosophy and methods of both competency testing and competency assessment are thus wholly different from the academic tradition of whole class, teacher-dominated instruction with testing at fixed points in the curriculum. Skills learned and not time spent are what drives the pace of instruction and assessment. Testing and assessment are not after the fact, external processes of inspection but integral parts of the process of education. In all of these respects, the traditions of testing and assessment in vocational education resemble what is being advocated elsewhere in the rest of education to replace standardized testing with alternative forms of assessment.

Whether the written methods of competency testing are any more or less reliable and more or less valid than the methods of assessment used in vocational education is impossible to say merely according to differences between the methodologies themselves. The critical issue with assessment is the comparability of judgments made across instructors and from program to program. This comparability concerns both the ratings or evaluations given for similar performances and the level of performance considered sufficiently high to meet the standards. With sufficient efforts to develop consistency in rating processes through procedures of training and group judging, and to utilize available techniques for statistically checking on the consistency of ratings, high levels of consistency in raters' judgments can be achieved in performance assessment.[10]

With written testing, the two most critical issues are the relevance of what can be measured to capabilities for performance, and the long-run effects of closed-form testing instruments on the content and methods of instruction. Written forms of testing are generally thought to be best for measuring factual knowledge and certain forms of reasoning. The "know-how" and capabilities for more complex and extended performance that are critical in the workplace (and in life) can generally not be measured as well with written forms of testing. Written tests suffer from the fact that the knowledge and powers of reasoning that can be measured with the greatest internal consistency of individual test items, which is the necessary criterion of a sound written test, typically do not include some of the most important capabilities for occupational preparation.

The crucial point is to not assume that methods of written testing are sound simply because of their written format and that methods of assess-

[10] The major source of variation in performance assessment has generally been found to be task variety rather than inconsistency in raters' judgments. See, for example, Richard Shavelson, "Generalizability Theory and Military Performance Measurements: Individual Performance," *Performance Assessment in the Work Place* (Washington, DC: National Academy Press, 1991), pp. 207-257.

ment are unsound simply because they involve elements of judgment in scoring performances. Much closer investigation of the consistency and relevance that are possible with different forms of testing and assessment, and how they are conducted in vocational education, is needed before conclusions can be drawn about which methods are best for what purposes.

It is important to point out that assessment in vocational education may include written testing as one of several methods of measurement employed. The emphasis in assessment is on using different methods of measurement for different kinds of knowledge and skill rather than heavily emphasizing written testing.

DIFFERENCES AMONG THE STATES IN POLICY ON TESTING AND ASSESSMENT

One major difference among the states is the emphasis that they place on written testing compared to assessment. Many advocates of competency-based methods of testing and assessment have encouraged the development of state resources for written competency testing for some time; certain states are much further along in the development of this capability than others. The Vocational-Technical Education Consortium of the States (V-TECS), described in chapter 4, has been one of the important results of this development.

Another important difference among the states is the extent to which local programs are required, or *mandated,* to follow the policy on testing and assessment or are only *encouraged* to follow it as a matter of state board policy, state administrative policy, or state legislation. A few states have no policy or program of testing and assessment for vocational skills but most do, and either require or encourage local programs to follow it.

Using these two dimensions of difference, the 50 states and the District of Columbia can be grouped into four different types of environments for testing and assessment:

1. *States that encourage written testing* for occupational skills in local programs, including competency-based or other forms of written testing. The encouragement of testing in most of these states is strong.
2. *States that mandate assessment* for occupational skills in local programs without specifying what methods should be used. These states typically encourage the use of multiple methods of assessment and testing in local programs.
3. *States that encourage assessment* for occupational skills in local programs without specifically encouraging one method over others. These states also typically support the use of multiple methods of assessment. The encouragement given to assessment in these states is generally not strong.
4. *States that have no specific policy or program* of encouragement or requirement for either the testing or assessment of occupational skills.

A fifth category of "mandated testing" turned out to include only two states in the 1992-93 school year so it was combined into the first category (states that encourage testing). One of these states is a small state that mandated one form of testing for occupational skills in 1992-93 and the other is New York, which has two types of statewide tests for occupational competencies.

The fact that only two states have mandatory testing shows one clear difference already between vocation education and the rest of elementary and secondary education. In academic education, most states have a large, statewide, mandatory program of testing for academic skills. For example, there were 39 states where a norm-referenced, standardized, written test of academic skills was administered statewide to all students in one or more subject areas (e.g.,

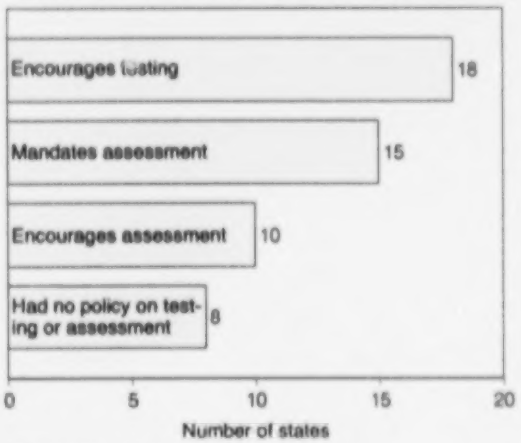
mathematics) in the 1992-93 school year.[11] In nearly all cases, these tests are administered at a certain grade level in those subject areas. No such mass, mandatory, statewide testing of all the students at a grade level or any other fixed point in time is conducted in vocational education. Furthermore, in many states this mass testing is administered by the state (or a contractor to the state) rather than local school districts. In vocational education, most of the testing and assessment is actually conducted or administered by staff of the local programs.

All 50 states and the District of Columbia were classified into the four categories defined above using data from the survey of states conducted by OTA. As shown in figure 3-1, 18 states are classified[12] as encouraging testing for occupational skills in 1992-93, 15 as mandating assessment for occupational skills, 10 as encouraging assessment of occupational skills, and 8 had effectively no policy on vocational testing or assessment.

The 18 states in the first category of encouraging testing enrolled about one-half of all high school students. The 15 states in the second category of mandating assessment enroll about one-quarter of all high school students. The last two categories each enroll about one-eighth of all high school students.

It is important to point out that the states classified as having no policy on testing or assessment for occupational skills may still have adopted performance standards and measures based on other kinds of information, such as rates of program completion or placement. Some of these states also chose to meet the performance standards requirements in the 1990 amendments by adopting a policy of allowing local programs to use their own performance standards and measures, which were then considered to be the state's performance standards and measures.

Figure 3-1: State Policies on Testing and Assessment for Occupational Skills (in 1992-93)[a]



[a] Fifty states plus the District of Columbia.

SOURCE: Office of Technology Assessment, 1993.

STATE RESOURCES FOR TESTING AND ASSESSMENT

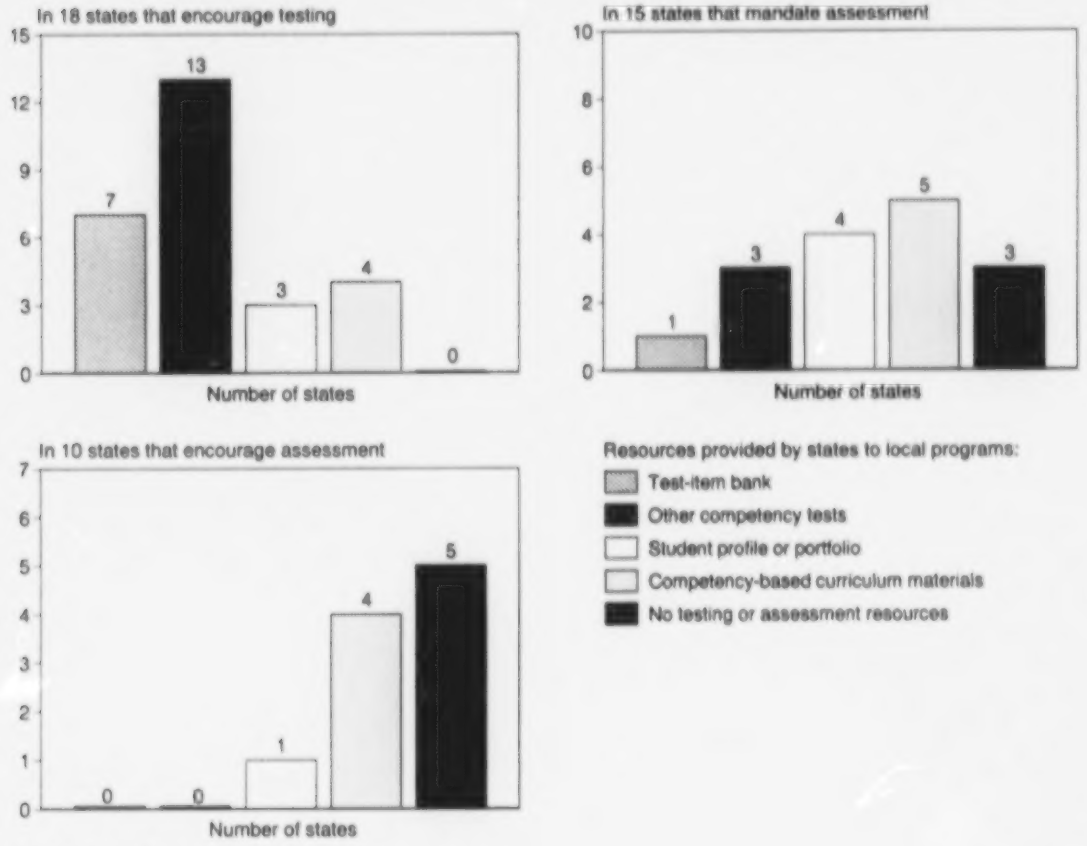
States encouraging testing for occupational competencies differ substantially in the extent of encouragement they provide, but there is no way of distinguishing among them using the data collected in the OTA survey. At one extreme are states like Oklahoma, where substantial investments have been made in competency testing and assessment in over 100 job-specific areas over a number of years. Oklahoma even has its own test development center, the Oklahoma Occupational Testing Center (OOTC). The state makes three kinds of instruments available to local programs:

1. tests for measuring the gains in the specific occupational competencies taught in specific courses,
2. tests for measuring the levels of competency achieved by students who complete sequences of courses, and

[11] North Central Regional Education Laboratory, *State Student Assessment Program Database, 1992-93* (Oak Brook, IL: Regional Policy Information Center, September 1993), table 3.7, p. 141.

[12] States are classified on the basis of their major component of testing and assessment for occupational skills.

Figure 3-2: State Resources for Testing or Assessment of Occupational Skills (in 1992-93)



SOURCE: Office of Technology Assessment, 1993.

3. profiles displaying those competencies that are given to students when leaving local programs to use in seeking employment.

At the other extreme is a state that only makes a single, state-developed test of "core occupational content" available to local districts to use as they wish.

Further differences among the four types of states can be illustrated by showing the kinds of resources for testing and assessment they provide or strongly recommend to local vocational programs. As shown in figure 3-2, four types of resources are provided:

1. *Competency test item banks.* Using these test item banks, local programs construct or write their own criterion-referenced competency tests reflecting the specific needs of their own programs of instruction.
2. Other forms of *competency tests.* These may be developed by the states themselves, purchased from vendors such as the National Occupational Competency Testing Institute, or obtained from industry groups for use in certifying students for particular jobs.
3. Student *profiles* or *portfolios.* These resources are used to provide students with a

means of reflecting on their own progress and communicating their accomplishments when seeking employment, further education, or for other purposes.

4. *Competency-based curriculum* materials. These materials frequently include testing instruments, assessment ideas, assessments, written testing, and/or checklists of competencies that are suitable for purposes of both teaching and conducting assessments.[13]

As indicated in figure 3-2, all 18 of the states that "encourage testing" either recommend the use of, or make available to local programs, a competency-based testing instrument or test item bank. Some of these states also provide the other two types of resources for assessment, but not nearly to the same extent as states that mandate assessment or encourage assessment. On the other hand, only four of the states that mandate assessment make resources for competency-based testing available to local programs.[14] These states apparently do so without specifically encouraging the use of testing over alternative forms of assessment. Both the states that encourage testing and those with mandatory assessment tend to place a strong emphasis on competency-based vocational instruction and embed assessment in the instructional process. These two types of states differ, however, in the extent to which they stress the need for written methods of testing, as opposed to more qualitative methods of assessment. In the states with mandatory assessment, local programs are required to conduct some form of assessment, but may use methods that best suit the philosophy and needs of their program.

The extent to which local programs are competency based differs among the four types of states. As shown in figure 3-3, states that mandate assessment tend to report higher percentages of competency-based programs than do states that encourage assessment or even those that encourage testing. This is true even among the subset of nine states where extensive test item banks for competency-based instruction have been developed. It may be that competency testing is being used in these states as a mechanism for forcing local programs to employ competency-based instruction and it is not working as well as the methods that are being used in states that are mandating assessment.

In fact, states with no policy of testing or assessment for occupational skills appear to have nearly the same incidence of competency-based instruction as states that encourage testing. This fourth group of states includes about equal numbers of very small states with no policy on testing or assessment for occupational skills, and larger ones where significant statewide educational reforms are under way that involve vocational education. (The number of states in the fourth category is very small, so that the apparent similarity with the first category may not be significant.)

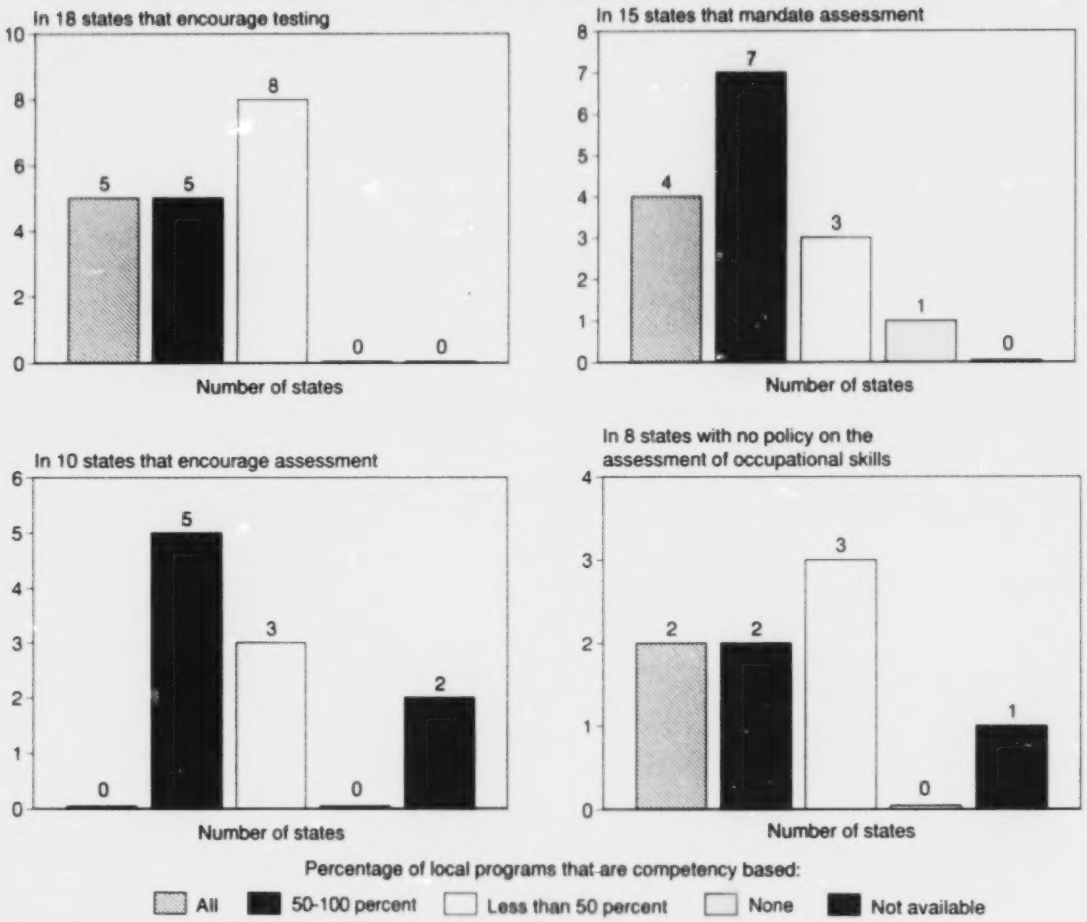
WHAT SKILLS ARE ASSESSED?

One of the most important aspects of the 1990 amendments is the substantive priorities for the reform of vocational education. Chief among these reforms is the integration of academic and vocational education, or combining the teaching of academic and vocational skills in the school curriculum. A second reform in the legislation is to broaden the occupational skills around which vocational programs are organized. This priority, which will be called organizing vocational education around *broad technical skills*, is less well developed in the legislation than the priority on integrating academic and vocational education, but is evident in the language of "applied technology" and teaching "all aspects of industry" in vocational programs that appears in the

[13] Because of the way that the questions were asked in the OTA survey, the data in figure 3-2 represent the main resources provided to local programs rather than all of the resources provided.

[14] The testing items in figure 3-2 include "Test-Item Banks" and "Other Competency Tests."

Figure 3-3: Use of Competency-Based Curricula in Local Programs (in 1992-93)



SOURCE: Office of Technology Assessment, 1993.

1990 amendments. It is also evident in the request from Congress for information from OTA about the availability of instruments for assessing the broad technical skills of vocational students. A third kind of reform that has been recommended by various outside commissions and studies is that preparation for work in vocational education and all kinds of training programs should be organized around the development of what in this report are called *generic workplace competencies* or skills.[15]

A major new development is the plan to create a National Skill Standards Board in the Goals 2000: Educate America legislation currently being considered by Congress. This calls for broadening the technical skills taught in vocational and other

[15] Anthony Carnevale et al., *Workplace Basics* (San Francisco, CA: Jossey Bass, 1990); and Secretary's Commission on Achieving Necessary Skills, *What Work Requires of Schools* (Washington, DC: U.S. Government Printing Office, June 1991).

kinds of work-related education and training programs. The board would do this by defining the skills required by industry in "broad occupational areas" of the economy and supporting the development of a national system for certifying the competence of individuals in those areas. The goals reflect recommendations that have been made by a number of outside commissions and studies.[16]

These proposals for the reform of vocational education (and other forms of skill training and education) provide a skills framework for describing the kinds of change that are occurring in the testing and assessment programs of states. Change in the types of skills being tested for or assessed by the states provides some indication of their priorities for the reform. The four types of skills are:

1. ***Vocational skills,*** which consist primarily of job-specific skills determined through job analysis and other tools of competency-based vocational education.
2. ***Academic skills,*** which, among the state vocational testing and assessment programs surveyed, are primarily reading, writing, and mathematics.
3. ***Generic workplace skills***, which are of two types: a) *employability skills*, such as work attitudes and knowledge of how to find a job; and b) *workplace competencies*, such as ability to work in teams, communicate information, solve problems, and manage resources.
4. ***Broad technical skills,*** which are the core skills and understandings of technology, information, and organization, and even history, needed to perform effectively within a range of occupations in an industry, such as manufacturing, finance, or hospitality.

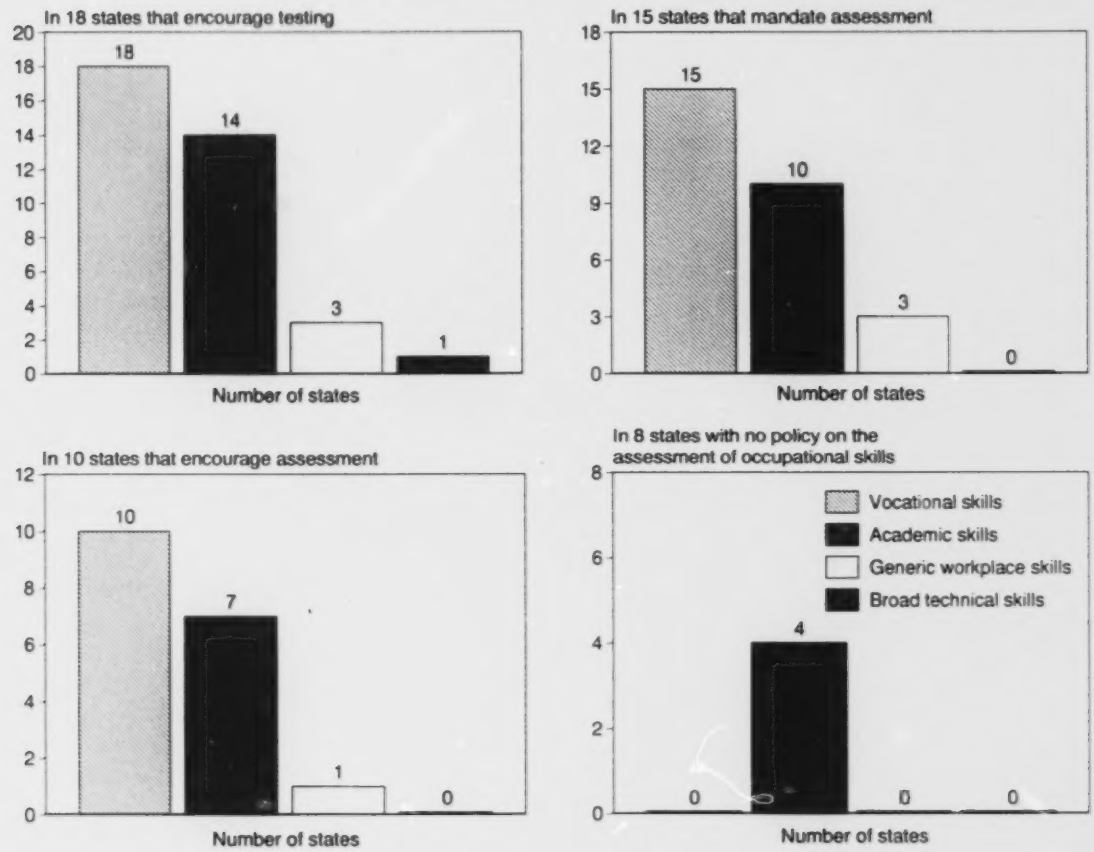
The first, third, and fourth of these types of skills together (vocational skills, generic workplace skills, and broad technical skills) will be called *occupational skills.*

The number of states with a component of testing and assessment for each of these four types of skills is shown in figure 3-4. The percentages of states having at least one component of testing or assessment are very similar for each type of skill—that is, the percentage of all states in the first category that test for or assess academic skills is about the same as the percentage of states in the second and third categories of states that test or assess for academic skills, and so forth. All 43 states in the first three categories of states have at least one component of testing or assessment for vocational skills. In the first three categories, 31 states have at least one component of testing or assessment for academic skills. Seven states in the first three categories have components of testing or assessment for generic workplace skills. Only one state in the first three categories supports testing for broad technical skills. In states with no policy of testing or assessment for occupational skills (the fourth category), only academic skills are assessed.

It is important to point out that the 1992-93 school year was actually the first year in which states were expected to operate their new performance standards under the 1990 amendments to the Perkins Act. This affects mainly the number of states with a component of testing or assessment for academic skills in subsequent years. If the baseline year of the OTA survey had been 1 year earlier, the number of states conducting testing or assessment for academic skills would in all likelihood have been much smaller than the 31 found in 1992-93, while the number conducting testing or assessment for vocational skills would

[16] Commission on the Skills of the American Workforce, *High Skills or Low Wages!* (Rochester, NY: National Center on Education and the Economy, 1991); and Gene Bottoms, *Redesigning and Refocusing High School Vocational Studies* (Atlanta, GA: Southern Regional Education Board, 1993).

Figure 3-4: Skills Included in Current State Policies on Testing and Assessment (in 1992-93)



SOURCE: Office of Technology Assessment, 1993.

have been about the same as shown in figure 3-4.[17] The reason is simply that most states essentially employed their existing policies of testing or assessment for vocational skills in implementing performance standards. Many fewer states initially had any capabilities in place for testing or assessing academic skills.[18] This indicates a substantial response by the states to the requirements of the 1990 amendments for performance standards.

[17] In a 1991 survey, the National Center for Research in Vocational Education found that 24 percent of states had previously used specific performance standards and measures for academic skills at some time in the past. E. Gareth Hoachlander, *Performance Measures and Standards for Vocational Education: 1991 Survey Results*, MDS-388 (Berkeley, CA: National Center for Research in Vocational Education, January 1992), figure 1.

[18] The number of states with a statewide academic testing program is large but the number of state vocational education agencies that were regularly obtaining data from those programs for vocational students was undoubtedly small.

WHERE IN THE CURRICULUM ARE THE SKILLS ASSESSED?

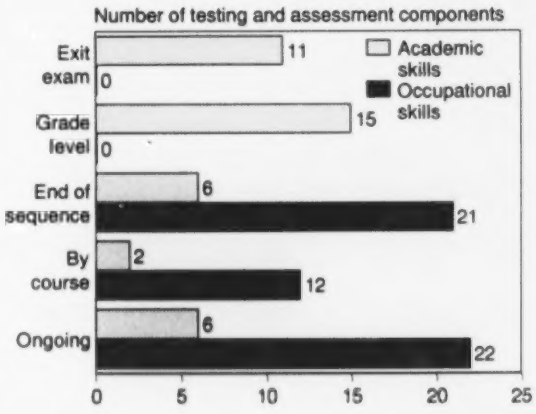
A second way of describing the testing and assessment policies of states is to ask where in the school curriculum the different types of skills are assessed and for what reasons. The answers shed some light on relationships between performance standards and reform.

The survey data collected by OTA show that the testing and assessment for academic skills included in state programs is highly separated from the testing and assessment for occupational skills. This separation appears to be undercutting the priority in the legislation on integrating academic and vocational education. The separation occurs because of major differences in the methods of testing and assessment being used for measuring academic and vocational skills.

As indicated in figure 3-5, most states have chosen to measure the academic skills of vocational students through obtaining test results from either their statewide academic testing program or their statewide minimum competency exit examination. Twenty-six of the 40 components of testing or assessment for academic skills in 1992-93 used these measures. These 26 components cover 25 of the 35 states where academic skills were measured in 1992-93, or 71 percent of all states. Fifteen of these 26 components use results from the regular statewide program of academic testing or assessment, and 11 use the minimum competency exit exam. The remaining 14 of the 40 components of testing or assessment for academic skills are much more closely tailored to students' patterns of enrollment in vocational and academic programs.

The problem with relying on exit exams and statewide academic tests is that they are administered centrally without any relation to when students enroll in vocational education, while testing and assessment for occupational skills is done locally and is closely tied to individual courses or the completion of vocational and academic courses.

Figure 3-5: Location of Testing and Assessment in the Secondary School Curriculum (in 1992-93)



NOTE: Occupational skills include vocational, generic workplace, and broad technical skills.

SOURCE: Office of Technology Assessment, 1993.

Furthermore, the bulk of this academic testing occurs in the 9th, 10th, and 11th grades, although some of the academic testing occurs in the 12th grade. (Some of it actually occurs as early as the 8th grade.) Minimum competency exit exams are taken for the first time in most states in the 9th or 10th grade and students keep taking them until they pass in order to graduate with a diploma.

In contrast, very little of the testing and assessment for vocational skills among the total of 55 components of testing and assessment for occupational skills shown in figure 3-5 occurs in the 9th or 10th grades. Most of the testing and assessment is conducted in the more advanced occupationally specific coursework that students typically take in the 11th and 12th grades. Only 6 of the 55 components shown cover vocational courses at the general introductory level, which are typically taken by students in the 9th or 10th grades. The remaining 50 components are for occupationally specific vocational education.

Another difference between academic and occupational approaches is that vocational skills measurement is much more closely related to specific courses and sequences of courses in the

vocational curriculum than is academic testing. As shown in table 3-1, in the states that encourage testing for vocational skills, the most frequent place in the curriculum for testing and assessment is at the end of a sequence of occupationally specific courses. This is the case by a substantial margin over all of the other possibilities in percentage terms. The second-most frequent pattern is to test or assess for skills in relation to specific vocational courses. These two uses of tests and assessment appear to reflect the close relationship between testing and competency-based vocational instruction that exists in states with strong policies of testing or assessment for vocational skills. The purpose of the testing and assessment at the end of a sequence of courses is generally to determine what students have learned.

The third basic pattern is ongoing testing and assessment for vocational skills, which means that local programs are encouraged to conduct testing and assessment as an integral part of instruction, with perhaps summative assessments or other forms of evaluation at certain points, without any specific guidelines on when the assessment should take place or for what specific purposes. Especially in the states with policies of mandatory assessment, this ongoing assessment is typically done within a framework of competency-based vocational education. The emphasis in these states is on encouraging local programs to consider multiple approaches to assessment and to develop their own programs.

The policies of testing and assessment for academic skills in these same (mandatory assessment) states are similarly ongoing, as shown in table 3-1. Few of these states have formal statewide testing programs like the states that are obtaining information about academic skills by grade level or from a minimum competency exit exam. How frequently local programs in these states are finding ways of linking academic and vocational assessment in their ongoing programs of testing and assessment is impossible to tell from the data OTA collected.

The potential for assessing academic and vocational skills together is much greater in the states where the policies of testing and assessment for both academic and vocational skills are highly localized than in states where the data are being obtained from a centralized, statewide academic testing program. *None* of the states in the two categories of "mandatory assessment" and "encourage assessment," where most of the ongoing testing and assessment occurs, indicated that a priority has been placed on the integration of academic and vocational skills in the practices of testing and assessment. They may be doing this but it is not possible to tell from the OTA survey.

There are only 14 components in 12 states where close course-level connections are apparently being drawn between academic and occupational skills assessment. These are shown in the top portion of columns three, four, and five of table 3-1. There are three basic patterns of test policy and use among these components. One is that a commercially available academic test, like the Test of Adult Basic Skills, is being directly administered to vocational students as they enter or complete specific vocational courses or sequences of courses. Gain scores are obtained in a few of these cases by correlating test scores from program or course completion with scores on the same test taken at an earlier point in the student's studies. This takes a substantial degree of coordination. The second pattern is states where academic skills have been incorporated into the lists of occupational competencies around which testing or assessment for occupational skills is being conducted. There is only one state (Oklahoma) where a sufficient priority has been accorded to this strategy to have indicated it on the questionnaire. However, pilot projects to accomplish this same strategy are currently under way in several other states. The third pattern is states where tests for academic skills that are contextualized to specific occupational areas have been developed. Arizona's Academic Test Item Bank (ATIB) is the clearest example of this. The ATIB is a

Table 3-1: Location of Testing and Assessment for Vocational and Academic Skills in the Secondary School Curriculum in 1992-93
(total number of testing or assessment components by type of state)

Type of skills and state	Location in the curriculum				
	Minimum competency exam	By grade level	Sequence of vocational courses	By vocational course	Ongoing
Academic skills:					
Testing encouraged	4	6	3	2	1
Mandatory assessment ...	2	5	1	0	3
Assessment encouraged .	3	2	1	0	2
No current policy	2	2	1	0	0
Total	11	15	6	2	6
Occupational skills:					
Testing encouraged	0	0	14	9	4
Mandatory assessment ...	0	0	6	2	10
Assessment encouraged .	0	0	1	1	8
No current policy	0	0	0	0	0
Total	0	0	21	12	22

NOTE: The total numbers of components for academic and occupational skills are 40 and 55, respectively.

SOURCE: Office of Technology Assessment, 1993.

resource bank of written test items for academic skills that are criterion-referenced to occupational skills and set in vocational problem situations.

Problems of integrating academic and vocational skills testing and assessment will be difficult to overcome in states that rely on test scores from their large-scale, statewide standardized testing programs for the assessment of academic skills. One problem with using this statewide data is that the *grade levels* where academic test score information is being collected are different from the grade levels where testing and assessment for occupational skills is being conducted. The fact that most of the academic testing is done before students take most of their vocational education suggests that academic testing could turn out to be used primarily as a screen for entrance into vocational programs, or as a prod to students' teachers in earlier grades, rather than as a means of verifying the academic skill attainments of vocational students when they leave their vocational programs or for actually improving the integration of academic and vocational learning.

There will also be difficult logistical problems in matching the statewide test data with the highly localized vocational testing and assessment on a student-by-student basis. Merging or comparing pre- and post-test score data for the same student from these two sources and relating any gains or losses observed to the very complex patterns of local enrollment in vocational courses and programs will be extremely difficult. The most likely use of the centralized academic data will be in coming to broad conclusions about the general levels of academic performance of students in vocational programs compared to other parts of the school curriculum. Whether the higher or lower scores of the vocational students are due to the vocational curriculum, the success or lack thereof in integrating academic and vocational instruction, or processes within schools of channeling students into different curriculum areas in the first place will be impossible to determine. The best solutions to these kinds of problems are likely to lie in the decentralizing of testing and assessment for academic skills to local programs and embedding it in instruction. The challenge will be to find a way of doing this while maintaining a sufficient degree of comparability across programs.

Why most of the states have chosen to import test data from their large-scale, statewide academic testing programs is reasonably clear. The data provide the simplest and cheapest means of complying with the requirements for standards and measures of performance "including academic skills" in the 1990 amendments. All other solutions are complicated and more difficult to implement.

WHY ARE THE SKILLS ASSESSED?

Another issue is how states are using information from their programs of testing and assessment to meet performance standards requirements. These could include using the results in making decisions at the state and local levels about the improvement of schools or programs, diagnosing student learning and modifying instruction at the classroom level on a regular basis, monitoring student attainment of skills in courses or programs, and certifying students' capabilities. For performance standards, the issue is whether the information is actually being used to make decisions about programs, and not just being collected in order to comply with the requirements.

The purpose of the performance standards is to make decisions about local programs, but it is possible to simply collect information about local programs and compute whether they are meeting the standards without doing anything with the results. In the OTA survey, the state respondents were asked to indicate all of the purposes for which each of the separate components of their testing and assessment programs were being used at the state and local levels. For each component of their program of testing and assessment, state respondents were asked to check off on the questionnaire which of the following uses were being made of the information:

- To satisfy requirements for accountability, *not including* accountability or performance standards under the Perkins Act.
- To satisfy requirements for accountability *including* accountability or performance standards under the Perkins Act.
- To make decisions about the improvement of programs, courses, or schools at the state or local levels.
- To assess students for program or course completion, certification, or progress in learning.

Respondents were instructed to check all uses that occur at the state and local levels with some frequency. They checked an average of 2.07 uses per testing and assessment component for vocational skills and 1.98 uses per component for academic skills.

One important finding from this question is that the patterns of use are opposite for academic and occupational skills. For academic skills, the most frequent use of information resulting from testing and assessment at the state and local levels is compliance with the requirements of the 1990 amendments. The most frequent use of information about occupational skills is monitoring the skill attainments of individual students or certifying student accomplishments. This information also informs teachers on the effectiveness of instruction.

The second main finding is that the least likely use of the information for both academic and occupational skills is making decisions about the improvement of programs, schools, or courses. This is potentially a problem since performance standards are supposed to be used in reviewing the effectiveness of local programs.

Further indication of the seriousness of the problem is the very low frequency with which information from testing and assessment is being used at the state and local level for reasons of accountability other than meeting the Perkins requirements. Information from testing or assessment is apparently not a very important basis for accountability at the state and local levels. This means that the test-based forms of accountability in the 1990 amendments are substantially differ-

ent from whatever the principal mechanisms of accountability currently are. Using information from testing or assessment for purposes of accountability and making decisions about the improvement of programs will be relatively new for most states.

States do have considerable experience in collecting data on rates of program completion or followup rates of employment and earnings. Whether this information is merely collected or actually used for purposes of improving programs or making decisions about them is the important question.

There are several possible reasons for the low rates at which information from testing and assessment is being used for program improvement. One is that states were not required under the legislation to link their performance standards with the review of local programs until the 1993-94 school year. Perhaps this year the frequency with which information from testing and assessment is used for program improvement will increase substantially. Another reason may be that states simply do not have enough experience yet to know precisely how information from testing and assessment can be used for making decisions about programs.

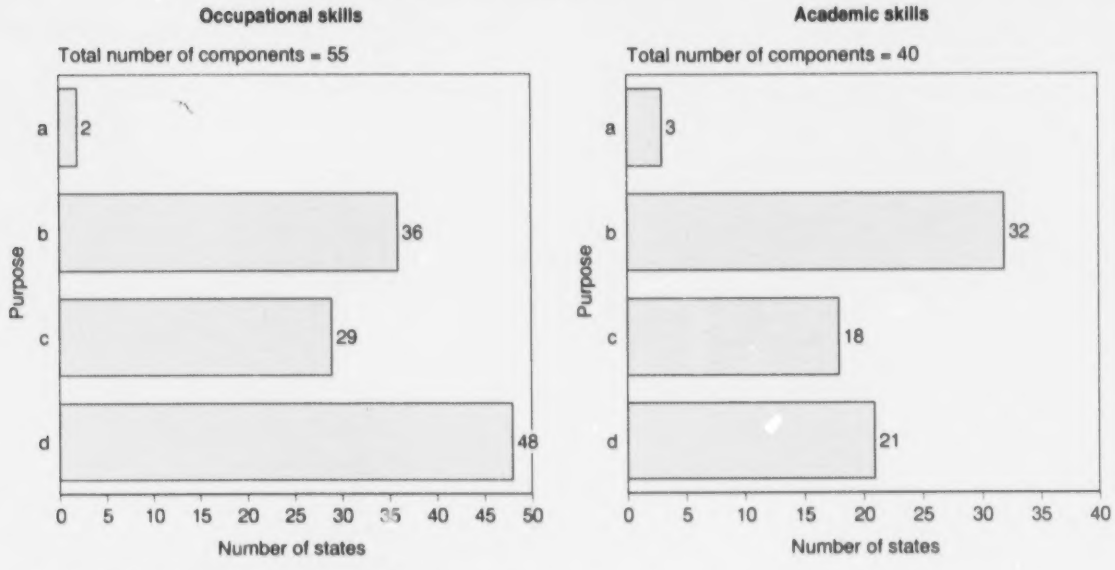
However, it may also be that the testing and assessment information simply turn out not to be very useful for making decisions about programs. The issues in appropriately using information from testing and assessment for making high-stakes decisions about programs are complex and difficult. Data are easy to come by. Information that really means something and measures accurately the outcomes of programs is much more difficult to obtain. Accurate information is even more difficult to obtain when the consequences of providing it to a state authority may be untoward. The examples given in the previous section about coordinating academic and occupational test information from different sources indicate how complicated and difficult the careful use of information from testing and assessment will be in many cases.

It is crucially important to recognize that the data presented in figure 3-6 can only be interpreted qualitatively because of the way that questions were asked in the OTA survey. For example, the data cannot be interpreted to say that information from testing and assessment for occupational skills is used "x" percent more often for working with students than is information about the academic skills of students. The only comparison that can be made from the data presented in figure 3-6 is that testing and assessment are used more frequently for some purposes than others, as discussed above, within the two categories of skills. In order to measure use in absolute terms, scales for recording perceptions of the importance of different uses of information from testing and assessment and combining them would have to be developed.

The data were checked to see if the basic profile shown in figure 3-6 differs among the four types of states. It does not. All four groups of states have basically the same profile.

Some differences exist in how information about occupational skills is used depending on where in the curriculum it is generated. As shown in table 3-2, the frequency of different uses is about the same for all three ways in which testing or assessment information is generated in the curriculum (at the end of sequences of courses, by course, or through ongoing assessment). The main exception to this is that information from ongoing assessment is significantly more likely to be used for purposes of student assessment than the information generated through the other two modes. Basically, the ongoing form of testing and assessment appears to be more student oriented than the two other forms of assessment. As will be shown in the next section, ongoing testing and assessment also involves less use of tests and more use of assessments than testing and assessment that is related to courses or sequences of courses. In short, testing appears to be oriented to accountability and assessment to student learning.

Figure 3-6: Purposes of Testing and Assessment in 1992-93



KEY: a = Used for accountability, not including Perkins
b = Used for accountability, including Perkins
c = Used for program, course, or school improvement
d = Used for the assessment of students for program or course completion, certification, or monitoring of learning progress

SOURCE: Office of Technology Assessment, 1993.

WHAT FORM OF MEASUREMENT IS USED?

The extent of reliance on written forms of testing as opposed to assessment differs greatly among the states, as shown in figure 3-7. Nearly all of the testing and assessment for academic skills encouraged or required by the states is done using standardized, written tests; only 7 out of the 40 components of testing or assessment for academic skills involve the use of instruments or methods other than standardized tests. In most of these cases, the standardized tests are developed by the states themselves, although some are commercially developed. The seven components of testing or assessment for academic skills are wide ranging. In Arizona, the vocational program will use results from the assessments for academic skills that each local school district is required to conduct under state policy. Some of these districts employ various forms of performance assessment, while others rely on written tests. Arizona has an item bank for assessing academic skills in the context of vocational skills. The vocational program in California will use results from their large, new, statewide system of performance assessment in the subject areas of reading, writing, mathematics, science, social studies, and history. In the four other states using assessment for academic skills, the state has turned directly to their local vocational program and expects them to develop means of conducting assessments for academic skills.

For occupational skills, the differences among the states are much greater. As shown in figure 3-7, the proportion of components that involve written testing is much higher in the 18 states that encourage testing for occupational skills than in the states that mandate or encourage assessment. In fact, 20 of the total of 25 existing components of testing or assessment for occupational skills in

Table 3-2: How the Results From Testing and Assessment for Occupational Skills are Used Depending on Where in the Curriculum They are Conducted (total number of uses)

Location of the testing or assessment in the vocational curriculum	Uses of the resulting testing and assessment information about occupational skills			
	For accountability purposes, not including Perkins	For accountability purposes including Perkins	For program, school, or course improvement	For student assessment and/or certification
At the end of sequences of vocational courses	1	15	12	16
During and/or at the end of individual courses	1	8	7	10
Ongoing in the vocational curriculum	0	13	10	22
Total	2	36	29	48

SOURCE: Office of Technology Assessment, 1993.

these 18 states involve written testing and only 5 others involve assessment. Conversely, only 3 of the total of 27 components in states that encourage or mandate assessment involve some form of written testing. The other 24 components involve some form of assessment. Generally, the model of assessment in these states is:

1. sets of core occupational competencies covering as many as 50 or more occupationally specific job areas are adopted by the state;
2. local programs are either encouraged or required to utilize these sets of competencies in organizing instruction and conducting assessments;
3. local programs are also typically encouraged or required to provide students with a profile of their competencies using the same competency frameworks; and
4. assessment tied to specific courses, program completion, or ongoing assessment is strongly encouraged or required.

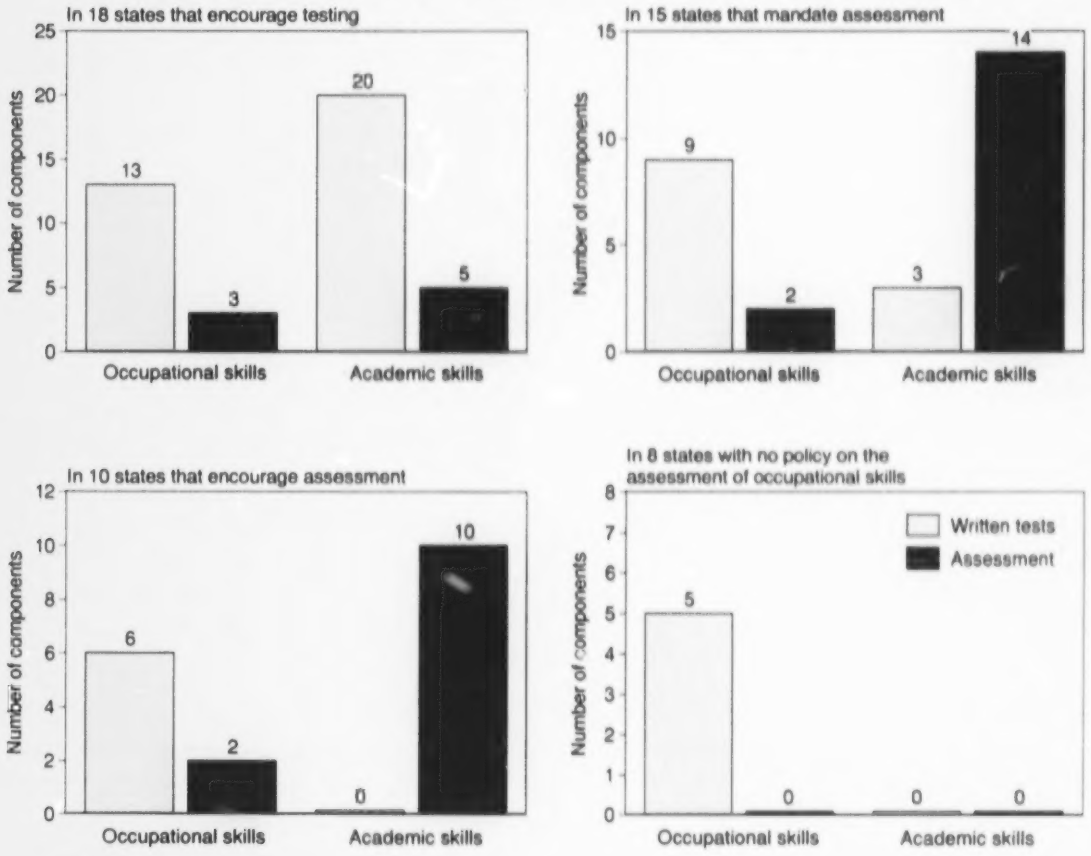
Especially in states that mandate assessment, local programs are furthermore required to:

5. submit evidence to the state confirming that the required assessments have been conducted and describing how they were conducted.

In the information about state programs that OTA obtained, only one state volunteered that local programs are annually reviewed using information from testing and assessment. (However, OTA did not specifically ask this question.)

While the information in figure 3-7 and data reported above distinguish sharply between written testing and assessment, it needs to be underscored that the line of separation between the two is blurred. Even in states where competency testing for occupational skills is encouraged, the main component of testing for occupational skills typically includes elements of both testing and assessment. This is also true for the major vendors of occupational tests, as described in the next chapter. In short, the components of "testing" that are shown in figure 3-7 frequently include some elements of assessment. Furthermore, states that encourage testing in their main component of measurement for occupational skills may also have other components primarily oriented to assessment. One example of this would be several states with both a policy of encouraging competency testing in their main component of testing or assessment for occupational skills, and providing students with profiles or portfolios. In fact, in vocational education there tends to be a substantial emphasis in most states on using multiple forms of assessment.

Figure 3-7: Methods of Testing and Assessment Employed by the States (in 1992-93)



SOURCE: Office of Technology Assessment, 1993.

The experience in most states where written forms of testing for occupational skills is encouraged is that the written tests end up being used to a greater extent than methods of assessment in measuring occupational skills. Once started, their use frequently tends to grow disproportionately compared to assessment. Even V-TECS, which is one of the key organizations involved in developing competency-based testing for vocational education, has performance items for only two-thirds of the 35 item banks where both competency lists and testing items are available, and the performance items comprise only a small proportion of each item bank.[19]

HOW ARE PRACTICES OF TESTING AND ASSESSMENT CHANGING?

The next question is how testing and assessment in vocational education are changing in response to the 1990 amendments. The most likely direction of change is that the performance

[19] Brenda Hattaway, assistant director, V-TECS, personal communication, Sept. 30, 1993.

standards that states have adopted will require expansion of testing and assessment.

OTA asked the states to describe their plans for the expansion, contraction, or continuation of their existing components of written testing and assessment for both academic and occupational skills over the 3 years between 1992-93 and 1995-96, and what new components of written testing or assessment for academic and occupational skills will be added. Questions were asked in some detail about changes in the skills to be assessed, the populations and programs to be tested or measured, and even the extent to which their plans for expansion are responding to the requirements for performance standards in the Perkins amendments. Questions were structured to determine if the plans are definite or are only being considered. Most states responded that their plans for expansion are definite.

The results show that a substantial increase in testing and assessment will occur in vocational education by 1995-96 (see figure 3-8). Most of the increase will occur in the expansion of existing components of testing and assessment rather than the creation of new ones, but some new components will also be added. In 1992-93, there were a total of 92 components of testing and assessment for academic and occupational skills among the states. Of those existing components, 48 are slated for expansion by 1995 and 20 new ones will be added for a total of 109 components by 1995.[20] The remaining 41 components will stay the same or nearly same through 1995.

States indicated in their questionnaire responses that the requirements of the Perkins Act for performance standards have been an important if not the deciding factor in over 80 percent of all these planned expansions of testing and assessment for occupational skills and over 70 percent of the planned expansions for academic skills.

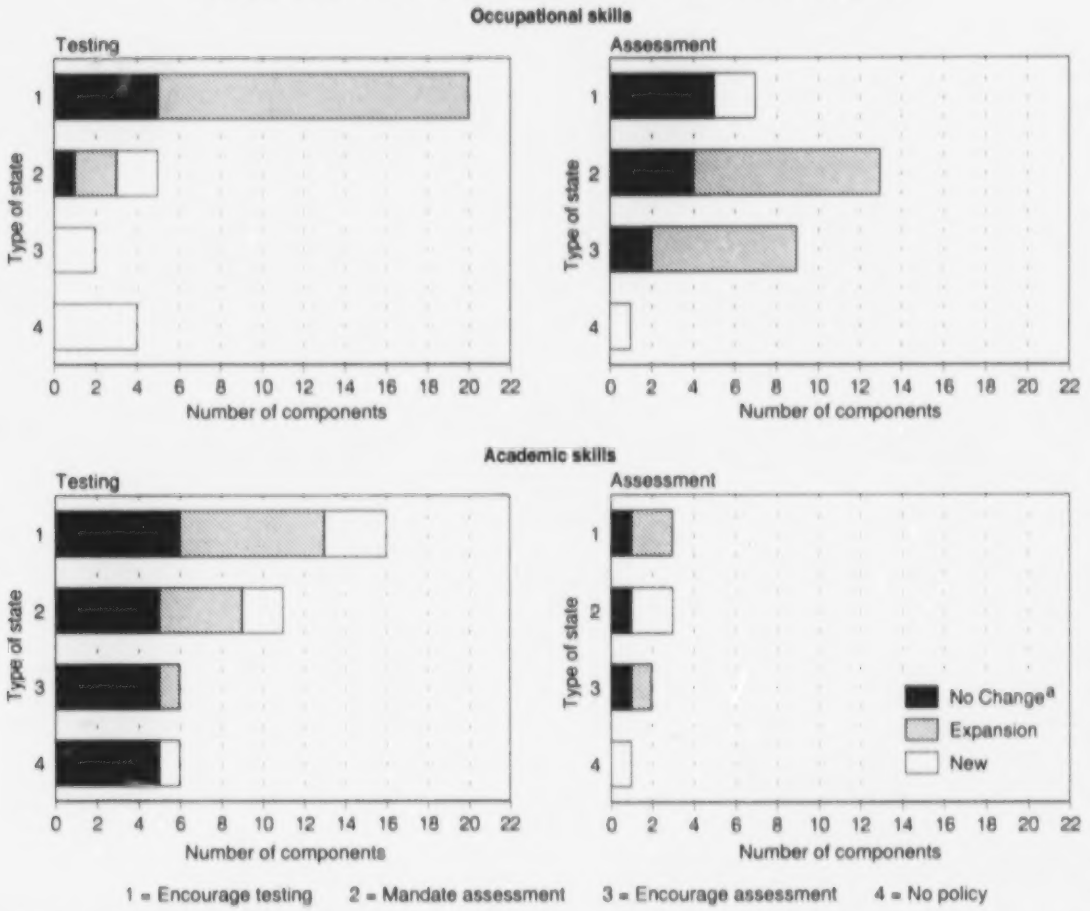
States' plans for the expansion of their policies of testing and assessment through 1995 cover a wide range. Some states are planning only minor additions or modifications to existing programs. Others involve the implementation of comprehensive new systems.

Ohio, for example, will implement an ambitious expansion of their 28-year-old Ohio Vocational Competency Assessment Program to include: a) three subtests from the new Work Keys system, which were being piloted in the state during 1992-93 (Work Keys is described in chapter 4); b) new or revised competency tests for both occupationally specific and employability skills in over 60 occupational areas; c) and a shift in policy from simply making competency tests available to local programs to strongly encouraging their use in local programs. In effect, a comprehensive statewide system of competency testing for vocational and generic workplace skills is being implemented. Ohio will also continue its longstanding statewide program of performance competitions for students.

Texas plans to add student portfolios to their currently required profiles of student competencies and to launch a new performance-based assessment program for generic workplace competencies. South Carolina plans to expand the number of occupational areas where test banks are available beyond the 41 areas where they are currently in use. Iowa plans to stiffen their requirements for assessing student competencies in local programs. West Virginia plans to add student profiles to their competency testing program in the 1993-94 school year and make the requirement mandatory for local programs. Kansas will incorporate generic workplace skills into their profiles of student competencies after piloting them this year. At the other end of the spectrum, there are several states where the only plans for expansion are to increase the number of

[20] Three of these new components will be existing components that change from assessment to testing. These three components are classified as "new" in figures 3-8 and 3-9. This explains why the 20 new components added to the 92 components that existed in 1992-93 increases the total number of components in 1995-96 to only 109. Two of these three "new" components are occupational and one is academic.

Figure 3-8: Planned Expansion of Testing v. Assessment by 1995-96



[a] One of these components will contract.

SOURCE: Office of Technology Assessment, 1993.

occupational or job areas covered in their existing program.

The most significant shift that will occur in the near future is from assessment for occupational skills to written testing. There will be eight new components of written testing for occupational skills among the states and three new components of assessment. However, two of the existing components of assessment for occupational skills will disappear for a net gain of only one new component of assessment. All eight of these new components of written testing will be in states that are currently in the category of either mandating or encouraging assessment, or do not have a policy on testing and assessment for occupational skills.

This means that the number of states in the category of encouraging testing for occupational skills will increase from 18 to 26. Those in the other three categories of mandating or encourag-

ing assessment, or not having a policy on testing and assessment, will decline from 33 states to 25 states.

Three-quarters of the existing components of written testing in states that encourage testing for occupational skills are also slated for expansion, compared to just over 50 percent of the existing components of assessment (17 of 23 components of testing compared to 16 of 29 existing components of assessment).[21]

The net effect of all these changes will be to shift the balance between testing and assessment in state policy from 23 components of testing for occupational skills and 29 for assessment in 1992-93 to 31 components of testing and 30 of assessment by 1995.

The major issue is what the long-term effects of this shift toward written methods of testing are likely to be on the character and content of vocational education. Programs of vocational education and training seem to be one of the last places where written testing should be preferred over methods of performance assessment. For example, methods of work sample testing have consistently been found to be more directly related to competence on the job than competency testing.[22] The long-term effects of written testing could be to shift teaching and learning in vocational education away from the development of more complex skills toward a greater emphasis on acquiring factual knowledge of various kinds, and away from constructive forms of student learning toward teacher-dominated modes of instruction. The effects on the character of vocational education could be profound.

There are several factors driving this shift toward increased written testing. One is the lower cost of written testing compared to performance assessment in time, materials, and instructor effort. This tends to produce strong incentives for the growth of written testing and the decline of assessment. Another is the widespread perception that written testing is more reliable and fair than methods of assessment where the judgments of instructors or evaluators are generally involved. In fact, with sufficient training and moderation of judgments, differences in rating among raters presents no problem.[23] The third is that the kinds of factual knowledge and information that are most readily tested tend to favor lecture methods of instruction over active methods of learning and hands-on laboratory and field work. This will tend to drive instruction in vocational education toward teacher lecturing and away from hands-on learning. For example, factual information about building codes and handbook information about the sizes and strengths of materials is much easier to incorporate into a written test and easier to teach with lecture methods than are concepts of designing and installing an electrical system for a house. Both are important but the issue is the balance between the two.

Unfortunately, there is very little good research on the long-term effects of written testing compared to assessment on the character and content of instruction in active learning environments like vocational education, so firm conclusions on this issue cannot be drawn. One of the best known studies is an article on training in the military, which gives one example of the dramatic effects that can occur. In the example, the substitution of performance assessment for written testing pro-

---

[21] Figure 3-8 shows 11 components of assessment that will not change and 16 that will expand. In addition, there are the 2 components of assessment that will be converted to written testing for a total of 29 components.

[22] Alexandra Wigdor and Bert F. Green, Jr. (eds.), *Performance Assessment in the Workplace, Vol. I* (Washington, DC: National Academy Press, 1991); and John Hunter, "Causal Analysis, Cognitive Ability, Job Knowledge, Job Performance, and Supervisor Ratings," *Performance Measure and Theory*, S. Lundy (ed.) (Hillsdale, NJ: Lawrence Erlbaum, 1983).

[23] Shavelson, op. cit., footnote 10.

duced many changes.[24] When written testing was stopped and students were assessed instead on performance, classroom lecturing on topics such as the muzzle velocity of guns and the functions of breech block locking bolts greatly diminished, desks were removed from the training rooms, equipment that the students were learning how to repair was brought in, and students began to spend much more of their time in class actually repairing equipment. The instructors complained that the new tests were too hard and they were right. The students' scores were very low on the performance tests. But their scores increased rapidly with each successive class of students in the program. Performance on the new assessments soon became the best predictor of grades in the class and correlations of the written tests with grades dropped precipitously. The only thing that was changed was the substitution of performance assessment for the written tests.

There is also a considerable amount of research showing that the imposition of high-stakes, standardized testing for basic skills on instruction in regular education tends to narrow the range of skills taught and learned by students.[25]

The irony is that written methods of testing appear to be expanding in vocational education at the very time that questions are being raised in the rest of education about the effectiveness of standardized testing, and a great deal of experimentation is under way with methods of assessment, student profiles, and the like.

There should be no mistake about what is being said here. Much of the expansion of testing and assessment that is now occurring in vocational education is in areas of assessment. A good example of one of the major new efforts in vocational assessment is the California Career-Technical Assessment Project (C-TAP—see box 3-A). C-TAP, which is in the process of being developed and tested, is a comprehensive system of assessment procedures and student portfolios organized in part around newly defined broad technical and occupational skills. But, in fact, this new system of assessment is only one of three new components of assessment planned for vocational education. And, meanwhile, two other new components of assessment in other states will be transformed into programs of written testing. The greatest expansion overall is occurring in methods of written testing.

WHAT SKILLS WILL BE ASSESSED?

By 1995, states are planning considerable change in the skills that will be measured. This change has implications for reform in vocational education.

How the mix of skills assessed will change is shown in figure 3-9 for the four basic categories of skills that have been defined: vocational skills, generic workplace competencies, broad technical skills, and academic skills. As in the previous figure, the pattern is basically one of growth. There will be expansion of the skill areas of testing and assessment.

Before discussing figure 3-9 in detail, it should be noted that change in testing and assessment is represented in both figures 3-8 and 3-9 by the components and subcomponents of states' policies of testing and assessment that are new, expanding, or staying the same. The data presented provide a proxy for more direct measures of the change that is occurring in testing and assessment at the local level, such as in the numbers of students who are being tested or assessed and the amount of time they are spending in testing or assessment. (Data like this would be extremely difficult to obtain.) The expansion of states' policies frequently involves, for example,

[24] Norman Frederiksen, "The Real Bias in Testing: Influences on Teaching and Learning," *American Psychologist*, vol. 39, No. 3, March 1984, pp. 193-202.

[25] George Madaus, "The Influence of Testing on Curriculum," *Critical Issues in Curriculum*, 87th Yearbook of the National Society for the Study of Education, Laurel N. Tanner (ed.) (Chicago, IL: University of Chicago Press, 1988), pp. 83-121.

Box 3-A: Career-Technical Assessment Project (C-TAP)

For the California State Department of Education, the Far West Laboratory for Education Research and Development is currently in the process of developing, pilot testing, and validating a statewide system of authentic student assessment for vocational-technical programs in the state. The system will be an integral part of the Career Vocational Division's occupational clusters strategy for vocation-technical education. The plan is to implement the system in high schools, adult education programs, and regional occupation programs/centers throughout the state.

Student mastery of core occupational content, as well as career skills and appropriate academic standards, will be assessed in an integrated way. The system involves the adoption of three kinds of standards: 1) a "framework" of occupational content standards in each of the cluster areas, 2) a series of so-called "career performance standards," and 3) standards for academic skills that are needed in the workplace and underlie the career performance standards. Two examples of these occupational clusters are animal science in agriculture and computer science and information processing in business. Career performance standards will be set in the areas of personal skills of attitude and time management, interpersonal skills, thinking and problem-solving skills, communication skills, occupational safety, employment literacy, and technology literacy.

Students will be certified as "work ready" for an occupational cluster. Student certification will signify that students have the knowledge, skills, and abilities to succeed in an occupational field; the system of assessment is primarily intended to certify students for further learning in an occupational area or an academic program of study. The curriculum standards may be changed in California so that all students will be certified as ready for work.

Two different kinds of assessments are being employed in the C-TAP system—cumulative and administered.

Cumulative Assessments

The cumulative assessments provide each student with a record of their accomplishments and levels of achievement in high school to use in seeking employment. Most of the assessment products are highly dependent on initiative taken by the students. The three main components of cumulative assessment are a Supervised Practical Experience (SPE), an assessment project, and a portfolio.

Supervised Practical Experience. Students who are enrolled in an occupational program requiring an SPE have their work supervisor complete an evaluation form rating their skills in the seven career performance areas. Additionally, students are rated on skills specific to their vocational program required "on the job." The completed SPE evaluation form is included in the portfolio.

Assessment Project. Students complete and present an approved assessment project during the course of their program. Projects involve either the planning and development of a tangible end product, a written description

increasing the number of occupational program areas where testing or assessment will be conducted, revising the testing instruments that will be available from the state and expanding their numbers, adding a new category of skills to the state's testing or assessment program, or making testing or assessment a requirement (mandatory). All such changes in policy would increase the impact on students. Most of the changes planned by states are of these kinds. Essentially then, figure 3-8 represents the change in the numbers of students who are being tested versus those who are being assessed, while figure 3-9 represents the growth that will occur in the number of students who will be assessed in different categories of skill.

It is important to point out a major difference between figure 3-8 and 3-9. Figure 3-8 shows the numbers of components of testing and assessment for academic and occupational skills. These are the same components as in figure 3-5, 3-6, and 3-7. In figure 3-9, each of the components of

and analysis of a significant "real world" experience in the student's field, or both. Project ideas will be developed jointly by the student and teacher according to specific guidelines provided by the state for each occupational cluster. Ideally, parents and business and industry people are involved in the development of the students' work. At the conclusion of the project, all students describe their learning experiences through an oral presentation.

Structured Portfolio. Students assemble a structured portfolio during the course of their high school program. The portfolio helps students organize and present a collection of their work for purposes of assessment and for presentation to prospective employers and/or advanced training institutions.

The portfolio can include a letter of introduction, a career development package, a research write-up, a completed SPE evaluation form, the assessment project results, and four or more work samples. The career development package will include a resumé, an employment or college application, and a letter of recommendation. The research write-up is accompanied by an original outline and at least one early draft that had been submitted for teacher review. The topics of the research write-up may be selected by the student from choices related to safety, social consequences, business practices, and future trends in the student's chosen field. The work samples are evidence of the student's best work and may be related to the student's certification project or SPE. Work samples are selected for inclusion by the students with guidance from the teacher according to criteria established by the state. Work samples may be represented by actual products, photographs, videotapes, performances, written descriptions, or any other reasonable and appropriate means.

Administered Assessments

The administered assessments are structured exercises that are given to students at a certain time in order to assess their capabilities for performance.

Project Presentation. Students present details of their assessment project to a review panel and respond to questions from the panel. The panel may consist of teachers, parents, students, and industry representatives. The student's presentation is evaluated according to specified criteria including oral presentation skills and ability to reflect on the project experience

Written Scenario. Students are presented with a written scenario representing a complex and realistic problem from their vocational area. Students have 45 minutes to respond in writing. They are judged on ability to apply content knowledge to address the problem presented in the scenario.

On-Demand Test. Students take an on-demand test focusing on the career performance standards. Enhanced multiple-choice items, including interpretative exercises and open-ended responses, are emphasized.

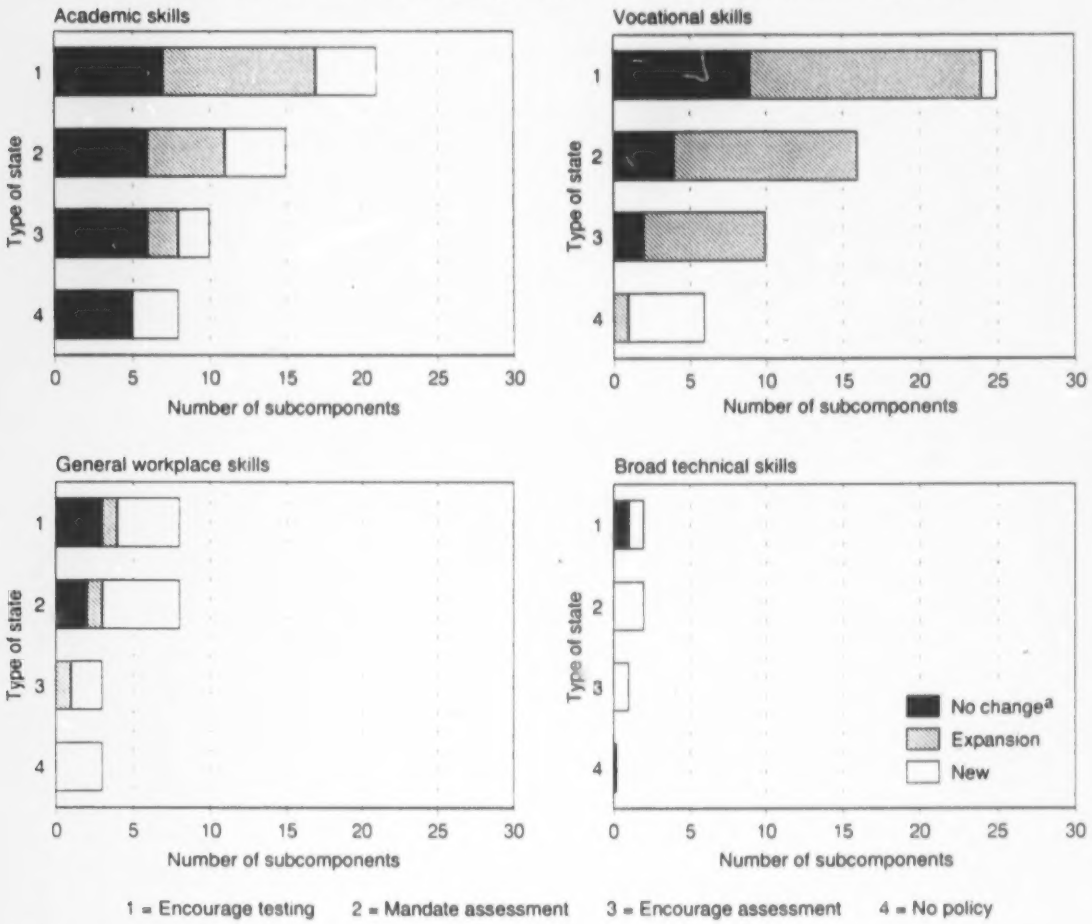
SOURCE: Adapted from Far West Laboratory for Educational Research and Development, *Career-Technical Assessment Project* (San Francisco, CA: Nov. 2, 1992).

occupational skills are broken into subcomponents according to the number of different occupational skills that are assessed. Many of the components for assessing occupational skills involve the measurement of two, or in a few cases all, of the occupational skills—that is, vocational, generic workplace, and/or broad technical skills. Each of these subcomponents are represented as a single count in figure 3-9 even though it may only be part of a more extensive component of testing or assessment for occupational skills. Included in figure 3-9 are 10 cases where both academic and occupational vocational skills will be assessed together by 1995-96.

Academic Skills

The most surprising result in figure 3-9 is that the slowest rate of expansion in state programs of testing and assessment will occur in the area of academic skills. State programs of testing and assessment for academic skills will *not change* in 24 of 41 components (59 percent) in which they

Figure 3-9: Skill Areas of Planned Expansion in Testing and Assessment by 1995-96



[a] One of these components of academic testing and assessment will contract.

SOURCE: Office of Technolgy Assessment, 1993.

were assessed in the 1992-93 school year.[26] The other 17 components (or 41 percent) will expand. Thirteen new components will also be added by 1995, which is a 27 percent increase in the total number of components by 1995. These new components will be mostly in states that had no means for testing or assessing academic skills in place for 1992-93. These percentages of expanding, unchanging, and new components of testing and assessment for academic skills will be approximately the same across all four types of states.

[26] One component of testing or assessment for academic skills will contract. To simplify figures 3-8 and 3-9, this component is included in the data for components where there will be "no change."

Vocational Skills

In contrast to the 41 percent rate of expansion for academic skills, testing and assessment for vocational skills will expand in 36 of the 51 subcomponents of state policy where they were assessed in 1992-93 school year, or at a rate of 70 percent of the existing subcomponents. In addition, testing or assessment for vocational skills will occur in six new components by 1995 for an increase of 12 percent.[27] The overall rate of expansion (including new subcomponents) for vocational skills will be even faster for generic workplace skills and broad technical skills, but the total amounts of testing or assessment for these two categories of skills in 1992-93 will still be considerably smaller than for vocational skills because of the smaller base. This means that among the four types of skills the largest total increase in testing and assessment will be for the conventional, occupationally specific skills of vocational education.

Integration of Academic and Vocational Education

This slow rate of growth for testing and assessment for academic skills is surprising for several reasons. One is that academic skills are a new area for testing and assessment in vocational education, and another is the strong indications in the 1990 amendments that *gains* in academic skills should be at least one of the two performance standards established by the states. Test score gains are generally much more difficult to measure than attainments at any point in time and especially in a statewide, mass testing program. The fact that many of the states are relying on test score data from their state-level minimum competency exit examinations or some other statewide test, which is administered at a certain grade level, makes it very hard to see how they will be able to show gains in students' academic skills over the course of their enrollment in vocational education. These difficulties indicate that state plans for expanding the capabilities for academic skills testing and assessment should be substantial.

Another apparent barrier to the integration of academic and vocational education is the substantial differences that exist between the methods of testing and assessment being used by states for academic skills compared to occupational skills, as discussed above. One of these differences is the heavy reliance on norm-referenced, standardized testing for academic skills compared to the criterion-referenced nature of testing and assessment for occupational skills. Another difference is that on the academic side, most of the testing and assessment is predominantly written, while on the occupational side, there is a mix of assessment and written testing. Furthermore, the testing for academic skills is typically centralized at the state level and conducted statewide on a mass basis at a given grade level, while on the occupational side the process is highly decentralized and tied to the local structure of courses and program completions.

A few statistics from the OTA survey tell this story of the divorce of testing and assessment for academic skills from occupational skills in stark terms. Only 10 of the 54 subcomponents of testing and assessment for academic skills planned by the states will be closely related to the testing and assessment for occupational skills at the local level—that is, in relation to the structure of courses or program completions, or in an ongoing program of assessment.[28] Various methods of doing this are being tried in the states involved but in only 5 of the 54 subcomponents are academic skills actually being tested for or assessed together with occupational skills in the context of the students' vocational and academic programs. In the other cases, a commercially

[27] While the rate of increase is only 12 percent, fewer states were without any capacity for testing or assessing vocational skills than academic skills. So the expected rate of new components of academic skills should be higher.

[28] The 10 subcomponents are not explicitly shown in figure 3-9.

developed academic test is being separately administered to students but in relation to their vocational course work.

The problems of reconciling academic and occupational test score and assessment information, and meaningfully interpreting it for purposes of monitoring or improving programs are formidable. Not paying close attention to these difficulties could lead to serious problems of misuse of test information and assessment results. If the content of the statewide academic tests does not reflect the academic content of the vocational or academic programs that are being integrated, then the resulting readings about the success of the efforts will be false. Also, the information from the academic tests may be in the form of "rank in class," as is frequently the case with standardized written testing, while the information about occupational skills will tend to come from assessments and again be much more competency specific.

Considering all of these difficulties, and the priority in the legislation on integrating academic and vocational education, it is surprising how little effort states are planning to devote to revising and expanding their policies of testing and assessment for academic skills compared to expanding their resources for measuring occupational skills. The ways in which testing and assessment for academic skills is separated from testing and assessment for occupational skills indicates that academic and vocational education are possibly being driven further apart rather than closer together. It is very hard to see how the testing and assessment information being produced by the states could easily be used for purposes of the highly localized, school-by-school, program-by-program, and teacher-by-teacher efforts that are required to integrate academic and vocational education. The slower expansion of testing and assessment for academic skills that is planned, together with the fact that academic and occupational skills assessment will be combined in only 10 cases, implies that the expansion of testing and assessment for vocational skills has higher priority for 1995 than the integration of academic and vocational education.

The inherent difficulties may explain a good part of the lack of plans for expanding academic skills assessment in the implementation of Perkins performance standards. Vocational educators may be understandably reluctant to initiate major coordination efforts with their statewide offices for academic testing, when this could involve substantial change in how testing and assessment is done and intrude in an area (the statewide academic testing program) where they have not been involved. The statewide office of academic testing may similarly be reluctant to coordinate their efforts with vocational education because it is unfamiliar territory and might involve substantial change in their approaches to testing and assessment. California's new portfolio-based system of testing and assessment for occupational skills, which includes some assessment of academic competencies, and where efforts will be made to coordinate with the new, statewide system of performance assessment, provides one model of how statewide assessment of academic and occupational skills can potentially be coordinated.

Generic Workplace Skills

A rapid expansion of testing and assessment for generic workplace skills is clearly in store. As shown in figure 3-9, the number of new subcomponents of testing or assessment for generic skills will exceed the number of existing components by nearly a factor of two for an increase of 175 percent. Most of the components of testing and assessment for generic workplace skills will be new in 1995. However, the large numerical size of this increase probably overstates the actual effects on learning and instruction in vocational education compared to effects of the planned expansion of testing and assessment for vocational skills, because generic workplace skills are the primary focus of testing or assessment in only 5 of the total of 22 cases. In the 17 other cases, the assessment

of generic workplace skills will be combined with the assessment of skills in one of the other categories of skills and be in a secondary role relative to those skills. Twelve of these 17 combinations are with vocational skills.

Broad Technical Skills

Even after the planned fivefold expansion, the smallest category of capacity for testing and assessment will be for broad technical skills. Several reasons for this can be surmised. One is the lack of clarity about the nature of these broad technical skills, and instruments and methods for assessing them. Another is that the reorganization of vocational curriculum around broad technical skills has not been the object of as much reform pressure as has been the concept of integrating academic and vocational education. However, the priority on broadening occupational skills appears to be growing. Still, the greatest total expansion of testing and assessment in absolute amount will occur in the traditional area of vocational skills, which tend to be very job specific (see figure 3-9). At least so far, performance standards appear to be working against the reform of broadening occupational skills.

The numbers of subcomponents where more than one category of skills are included is growing. For example, in figure 3-9, the total number of subcomponents shown for the 109 components of testing and assessment planned by the states for 1995 is 138. This is an average of 1.26 skill areas per component of testing or assessment. The comparable number for the 1992-93 school year was 99 skill areas for the 92 components of testing or assessment, or 1.08 skill areas per component.[29]

[29] In figure 3-8, state programs that involve the integration of testing and assessment for academic and occupational skills are treated as two separate components. Components of testing are "new" if they did not exist in 1992-93, or they will change by 1995 from either testing to assessment or vice versa. Subcomponents of testing or assessment for skill area are classified in the data as "new" only when the component is new. The addition of a new subcomponent of, for example, generic workplace skills to a program of testing or assessment for occupational skills is classified as a "new" aspect of an existing component.

Case Studies: Three Vendors of Tests for Occupational Competencies

4

Another perspective on the capacity for testing and assessment in vocational education is provided by describing three main vendors who supply states and local programs with resources for occupational testing. The vendors represent the mainstream of current practices in vocational education and one direction of future development—generic workplace skills.

The case studies include:

- the Work Keys System being developed by American College Testing (ACT);
- the Vocational-Technical Education Consortium of the States (V-TECS), which supports a system of test item banks for competency-based testing and assessment; and
- the Student Occupational Competency Achievement Testing (SOCAT) program of the National Occupational Competency Testing Institute (NOCTI).

The first assessment program is just now being developed and implemented. The latter two programs have been in operation for nearly 20 years.

WORK KEYS[1]

Origin of ACT and Work Keys

Work Keys is a system being developed by ACT for teaching and testing general workplace competencies and employability skills. The system is well along in development. When fully

[1] Joel West, Joy McClarity, Kate Ulmer Sottong, and Barry Mason, American College Testing, OTA interview, Jan. 28, 1993.

operational, as expected in early 1994, Work Keys will include the following four interrelated components for each of eight skill areas: 1) tests suitable for large-scale, high-stakes testing situations; 2) a job profiling component for analyzing the skill levels required in eight areas in real-life jobs; 3) instructional materials related to the skills assessed; and 4) a reporting service. Portions of all the components are available now.

Work Keys represents a broadening of mission for ACT, an independent, not-for-profit organization founded in 1959, which provides programs and services in the areas of college admissions, career and educational planning, financial aid, continuing education, professional certification, and licensure. ACT is best known for its testing program for college entrance. More than a million high school students take the ACT tests each year; college admissions officers use the scores in making admissions and placement decisions. In 1991, ACT decided to expand its services to encompass students bound directly for the workforce.

The Work Keys System

The broad goal of the Work Keys system, according to ACT, is to help strengthen the workplace skill competencies of all individuals who are bound for work or are already in jobs—not just vocational students. ACT hopes that schools will use Work Keys to help students see the connection between the skills acquired in school and those needed on the job.

The current design for Work Keys focuses on 12 different skills, each of which will eventually have its own separate test.[2] ACT identified these skills by surveying the literature on workplace skills and consulting with employers, teachers, and employment and training experts; the aim was to identify skills that are both measurable and teachable and that are viewed as important by employers and educators. When fully operational, Work Keys will enable test takers to evaluate their skills in a general way or compare their skill levels with those required by employers for specific jobs.

The four linked components of Work Keys can be summarized as follows:[3]

1. *Testing Component.* This will include at least 12 workplace skills tests or assessments that will be criterion-referenced and require written or other kinds of responses. The instruments will measure the level of competency demonstrated by the individual, or "how much" of the workplace skill they can demonstrate. The tests will be administered in a variety of formats such as multiple choice, constructed response, and computer adaptive. Necessary materials will range from paper and pencil to audiotapes and videotapes.
2. *Job-Profiling Component.* This component will enable a company to profile the competency or skill levels required for an employee to perform selected jobs successfully. The job profiling system was released in the fall of 1993.
3. *Instructional Support Component.* Instructional materials and accompanying reference guides will help learners and their teachers take steps to improve and broaden learner skills, so that people can obtain the jobs they want.
4. *Reporting and Recording Component.* A comprehensive recording and reporting service will provide informative reports on assessment results to students, teachers, and employers. For example, teachers may use the service to see how many of their students have strong workplace skills or evaluate instructional programs. Educators

[2] Although ACT plans to develop Work Keys tests for 12 skills, this number may change during the development process.

[3] The whole system is being implemented in phases and is currently operational for the initial skill areas; additional programs, services, and skill areas will be added over time.

and employers will be interested in comparing the test results to job profiles.

Purchasers will be able to buy and use any combination of the four components, although ACT will encourage users to view the system as a whole and to use all the parts for which they have a need.

The Work Keys Assessments

All of the Work Keys assessments aim to measure "work-related basic skills," with an emphasis on workplace applications of skills rather than academic applications. In addition, all of the assessments are criterion-referenced (not norm-referenced), meaning that an examinee is evaluated on his or her performance relative to the content and level of the test items and not the performance of other test takers.

Each test will include questions across a range of levels of difficulty, from four to six depending on the assessment. For each assessment, the range of levels reflects the skills required from the lowest level likely to be assessed to just below the level at which specialized training would be required. The levels are hierarchical. For example, an examinee who scores at the fourth level on the Applied Mathematics assessment should also be able to perform well on exercises at all levels below, but not levels above. Because the tests are criterion-referenced, a specific level on one assessment does not correspond to the same level on another assessment.

Six tests were released in September of 1993:

1. *Reading for Information.* This "... measures the examinee's ability to read and understand work-related instructions and policies. Reading selections and questions based on the actual demands of the workplace appear in the form of memos, bulletins, notices, letters, policy manuals, and governmental regulations." Questions on the test fall across five levels of difficulty, arranged from easiest to the most difficult, and are followed by multiple-choice questions.
2. *Listening and Writing.* This assessment measures listening and writing skills together, simulating how they are often used in the workplace. The examinee listens to audio-messages on tape, takes notes, and composes a written message. The written messages are scored in two different ways: a listening score measures accuracy and completeness of information, and a writing score addresses grammar.
3. *Applied Mathematics.* This measures the test taker's skill in setting up and solving work problems using "mathematical reasoning skills." Examinees are allowed to use calculators, just as in the workplace.
4. *Applied Technology.* This paper-and-pencil, multiple-choice test measures an individual's skill in solving technological problems, covering the basic principles of mechanics, electricity, fluid dynamics, and thermodynamics as they apply to machines and equipment found in the workplace. The test emphasizes skills in identifying and analyzing relevant parts of problems, evaluating potential solutions, selecting materials and solutions, and applying methods to novel challenges and circumstances.
5. *Teamwork.* This test measures an examinee's ability to "... choose behaviors and/or actions that simultaneously support team interrelationships and lead toward accomplishment of work tasks." Test takers will watch a videotape of teams of workers performing tasks and will be asked multiple-choice questions about the teamwork scenarios.
6. *Locating Information.* This multiple-choice test will measure the ability to use graphic documents to insert, locate, compare, and summarize information. The types of graphics used on the test include diagrams, blueprints, floor plans, tables, forms, graphs, and instrument gauges.

ACT plans to complete at least five additional tests in the next several years. One, *Observations*, is scheduled for release in September 1994. The others are currently in the design and early development phases.

- *Speaking*. This test will measure whether a person can convey a spoken message clearly.
- *Observation*. This test will measure a person's "watching skills"--the ability to learn from demonstrations.
- *Motivation*. This test will measure dependability and responsibility, and will focus on work habits that can be taught (rather than attributes of personality).
- *Learning*. This test will measure a person's skill at adjusting to changes in a job situation, such as those resulting from a new job or new technology.
- *Managing Resources*. This test will measure a person's skill in scheduling, budgeting, and allocating resources.

The Assessment Development Process

ACT is undertaking several steps to ensure that Work Keys will be responsive to the needs of employers and educators and that the assessments developed will be reliable, valid, and fair. Early in the development process, ACT created a consortium of six "charter states" with a demonstrated interest in new vocational assessments to give advice and help pilot the system.[4] No state is obliged, however, to use the Work Keys assessments once completed. Each charter state also has a Work Keys advisory panel, composed of two-thirds business and industry representatives and one-third educators. The advisory panels help ACT with the conception and development of Work Keys components, with prototype testing and pretesting, and with marketing the system. In addition, the panels are expected to help facilitate the use of the Work Keys assessments in their own states.

ACT follows a typical process of objective test development to ensure that test items and assessments meet high professional standards of reliability, content validity, and fairness. First, "constructs," or definitions, are identified and developed for each skill, with the help of two representatives from each charter state. Once constructs are developed, ACT attempts to define and describe a hierarchy of content levels for each skill to provide a set of criteria for test construction.

The second step is to draft items that correspond to each level of difficulty; ACT hires item writers to help with this process. All items must be written to the specifications developed by ACT and the advisory panels, and all items are edited by ACT staff.

The third phase is "prototype testing." In prototype testing, the draft items are administered to small samples of students and employees. ACT determines whether the draft items appear to correspond with the expected level of difficulty and satisfy the content criteria. Based on the findings, ACT rewrites and redevelops items for each level of difficulty.

Fourth, a large number of items are written, edited, and reviewed by experts for content validity and fairness. These items are pretested on a large sample of individuals. For the multiple-choice Work Keys tests, this sample consists of over 1,600 student and employee volunteers. The pretest results are analyzed for consistency and reliability. High-quality items meeting all content and statistical criteria are selected to produce the final tests.

Test Administration and Reporting

The Work Keys tests are designed to be administered by teachers or guidance counselors, in accordance with procedures proscribed by ACT. Currently, all tests completed by students

[4] The six charter states are Wisconsin, Tennessee, Michigan, Iowa, Oregon, and Ohio.

are returned to ACT for scoring; local onsite scoring will be available in the future.

For each test purchased, five different types of reports will be generated by ACT and sent back to the client.

- *Chart Essay Report.* This report provides some general descriptive information about how various groups of examinees scored on each test. It is organized around a standard set of questions such as: "Do the scores of males and females differ on the Reading for Information test?" For each question, a page of the report provides information about the percentages of examinees achieving different score levels in both bar and table form. The report is produced for each test administered.
- *Individual Reports.* Each of the examinees will receive a multipart report describing his or her performance on each test with suggestions of learning activities that might be undertaken to improve skills. The first part gives the examinee's scores on the assessment, a few tasks illustrating the levels of performance associated with the scores, and suggested learning activities. The second part contains the examinee's scores along with demographic and other information provided by the examinee during the testing process. The third part summarizes the information in a form suitable for attachment to job or school admission applications. The information in the report is designed to be used for course planning, career guidance, and individual goal setting.
- *Roster Report.* One copy of this report, which lists the name and four lines of information on each examinee, will be provided for each client order. The information will include assessment scores and demographic and job-related information (e.g., job-seeking and career choice information).
- *Vocational Information Reports.* This report is designed to be used in determining the career goals of a group of examinees and whether those goals match the occupational opportunities in a given region or city. It shows the percentages of examinees holding and expressing interest in particular jobs selected from a given list.
- *Local Items Report.* Clients may also include locally generated items in the Work Keys system. These questions are primarily intended to obtain information about an examinee's instructional experiences (e.g., in using calculators). This report tabulates the responses of examinees to those questions.

ACT is considering offering electronic resumes for interested individuals who take the Work Keys assessments. This resume would contain cumulative information on an individual's skill levels over time, and would be made available to prospective employers at the individual's request.

Job Profiling

The job profiling component of Work Keys will enable employers to identify the nature and level of work-related basic skills required for jobs in their companies. By following the Work Keys procedure, an employer will be able to determine the level of each of the 12 Work Keys skills required for every job profiled. Analysts trained and certified by ACT will conduct the job profiling. The profiling procedure is being developed in a joint effort with a number of companies.

Individuals who participate in Work Keys will also be able to develop their own skill profiles based on their assessment score reports. A student could then compare his or her personal profile to the job profile for their desired occupation. Both learners and employers will be able to see the extent to which an individual has the skills needed to qualify for a particular job. ACT is developing a database of job profiles (without employer identification) that can provide employers and educators with a general picture of skill requirements for different occupations.

The instructional component will consist of materials describing the Work Keys skills in greater detail and illustrating workplace applications of those skills. ACT may also offer workshops for educators and employees to discuss strategies for building the skills needed in the workplace.

Implementation of Work Keys

ACT began marketing the Work Keys system in early 1993. A number of states have decided to use or are considering use of the Work Keys system in various ways. Several states intend to use Work Keys for measuring "basic or more advanced academic skills" as required in the Perkins Act. Others may administer it as a pretest at the beginning of 11th grade and a post-test at the end of 12th grade to assess student gains in workplace skills. Still others may use Work Keys as a program completion examination at the end of grade 14, after the final years of an integrated secondary-postsecondary "tech-prep" program.

VOCATIONAL-TECHNICAL EDUCATION CONSORTIUM OF THE STATES[5]

Origin and Purpose of V-TECS

V-TECS is a consortium of 23 states with the goal of promoting competency-based vocational-technical education. Since its inception in 1973, V-TECS has been a unit within the Southern Association of Colleges and Schools, the main accrediting agency in the South. Full membership in V-TECS is limited to state agencies responsible for administering the Perkins Act programs. Member agencies provide proportional support for V-TECS administrative and product development costs. Associate membership is open to "... the military services, federal, state and local governmental agencies, international entities, and other organizations . . ." with demonstrated interest in performance-based education.[6]

V-TECS aims to accomplish its goal of promoting competency-based education through cooperative research and development (R&D) efforts in four main areas: 1) analyzing jobs; 2) organizing job-related information; 3) developing components for assessing student achievement; and 4) designing, developing, and/or acquiring instructional materials that link teaching with the skills required for jobs.

These four efforts are interrelated, with the first effort, occupational analysis, providing the foundation for the other three. The founders of V-TECS felt that the improvement of curricula for vocational-technical education should begin with occupational analysis—ascertaining the specific tasks and duties performed by workers in certain jobs and building a curriculum around them. The task of developing the occupational analysis was divided among participating states, with each state taking responsibility for certain occupations and sharing their findings with other members.

Assessment did not become a major focus of V-TECS until 1986, when banks of test items for states or other V-TECS members to use in constructing their own competency-based tests were developed to respond to the growing interest in better assessment and credentialing for vocational students.[7] As with curriculum, V-TECS members felt that assessment should be based on what students would be required to do in the occupations for which they were trained. Thus, every test item selected for the banks is tied to a specific task or duty in a specific job area.

---

[5] Information in this case study is based on Brenda Hattaway, assistant director, V-TECS, OTA interview, Jan. 12, 1993.

[6] The technical training commands of the Army, Air Force, Navy, and Marine Corps are associate members and support V-TECS through such means as sharing task lists with member states. Other federal agencies, such as the Bureau of Prisons and the International Labor Affairs Division of the U.S. Department of Labor, are associate members.

[7] According to Hattaway (op. cit., footnote 5), the main purpose of the tests has been to improve instruction.

Paralleling its four R&D efforts, V-TECS offers four main products and services:

1. *Analytical Tools.* First developed in 1973, these are called "catalogs" and exist for over 200 job areas.[8] For each job area, the catalog consists of lists of the duties and tasks involved, along with the tools, equipment, and work aids needed to perform them. Finally, for each of the duties and tasks there is a list of performance objectives and the steps that the worker must take in performing them.[9]
2. *Instructional Tools.* These, too, are organized by job area. For each job area, they include instructional worksheets, lists of instructional resources and activities, lists of enabling competencies, and lists of related academic skills. Development of instructional tools began in 1984, and they now exist for 66 job areas.
3. *Assessment Tools.* V-TECS has also developed or acquired banks of test items. Each of the test items is criterion-referenced to a specific duty and task. A test item bank is available for 35 of V-TECS's job areas. The banks include both written (matching or multiple-choice) and performance-based items.
4. *V-TECS DIRECT.* This is a software package designed for storing and retrieving the V-TECS materials that make up the other three components. As of fall 1992, catalogs and instructional tools for about 70 occupations were available on disk. All test item banks have been available on disk from their initial release.

Development Process for V-TECS Catalogs and Instructional Tools

V-TECS follows a multistage process in developing all of its catalogs and instructional tools. First, the consortium determines the priorities for job analyses,[10] new product development, and revisions by collecting statistical information from such sources as the U.S. Department of Labor and surveying the needs of member states.

Second, V-TECS identifies competency-based curriculum materials and other materials for each job area that it plans to analyze. These materials help V-TECS develop task lists and lists of tools, equipment, and work aids. V-TECS refines the lists by identifying a target population of workers in the particular occupational domain being analyzed. A sample of this population is interviewed and observed. The tasks are then organized under duty statements.

The task and duty lists and the tools, equipment, and work aids lists are then sent to a larger sample of the target population in survey form. Surveyed workers are asked whether they perform the tasks listed on the V-TECS lists and whether they use the tools, equipment, and aids listed. V-TECS analyzes the survey data to determine the percentage of incumbent workers that performs each task and uses the various tools, equipment, and work aids. Based on these percentages, final lists are developed.

V-TECS has recently taken steps to improve the development of task lists, so that they will be truly national in scope. In prior years the lists were validated by surveying people in only a few states. However, V-TECS has expanded its surveys to include industry people in many states.

---

[8] V-TECS estimates that its catalogs cover at least 90 percent of the vocational education program areas that exist in the nation.

[9] According to the *V-TECS Technical Reference Handbook:* "... each 'task' is a statement of measurable behavior which describes a meaningful unit of work having a definite beginning and ending point. A 'duty' is a broad descriptor under which similar tasks are organized." The performance objectives consist of a list or description of the conditions under which a task is performed, a description of the task to be performed, and a performance standard. This standard is "... an observable or measurable standard of performance deemed appropriate for successful completion of the task by incumbent workers."

[10] A job area may include one or more Dictionary of Occupational Title (DOT) job classifications.

Third, V-TECS selects a group of workers to serve on a writing team; the team is responsible for specifying performance standards and identifying the steps involved in performing each individual task in a duty area. Following this phase, subject matter experts are asked to identify the cognitive knowledge, psychomotor skills, and work-related behaviors that are critical to the performance of each of the tasks. The "enabling competencies" and "related academic skills" that eventually appear in the instructional tools are derived from these knowledge, skills, and behaviors.

The V-TECS Item Banks

All test items in the V-TECS banks are developed by member states in accordance with a standard V-TECS development model. Eleven states have contributed to the development of test item banks so far, with the V-TECS central office monitoring quality. If a member state wishes to have its own state-developed test item bank labeled as a V-TECS item bank, it must document that it has gone through the standard V-TECS development process. V-TECS staff also review and edit the state-developed test item banks, on occasion sending them back to the states for further editing.[11]

The V-TECS process for developing test items includes several steps:

- *Validating Task Lists, Performance Objectives, and Performance Steps.* The first step in developing test item banks is to review the continued validity of the task lists, performance standards, and performance steps contained in the catalog of the job area for which the test is to be developed. This is accomplished by having five to eight workers from different-sized companies and employers who are experts in the job area review the lists and make minor changes as they see appropriate.
- *Writing Test Items.* V-TECS selects a team of test item writers. The team must include instructors with recent work experience, the state technical coordinator or project director, workers from the domain area, and a V-TECS central office representative, when available. After receiving special training, the writing teams develop both written and performance items by reviewing the completed task analyses and using V-TECS guidelines to ensure that test items match the tasks.
- *Reviewing and Editing Test Items.* The test items are reviewed by four groups of experts and revised if necessary. The four groups include the writing team, test item construction experts, subject matter experts, and a sample of workers. V-TECS central office staff review the test items to make sure they match the duties and tasks of each of the occupations and to make sure they are formatted correctly. V-TECS staff then edit the test items and compile the test item bank.
- *Field Testing of Item Bank.* All test items are field tested in schools to check whether the items are clear, reliable, and free of sexual or racial bias and whether the directions are clear. Test administrators are asked to provide feedback on any difficulties encountered during the field test. Item response data from field testing are scored and analyzed for item difficulty, item discrimination, and distractor response.[12] These three forms of analysis help ensure that test items discrimi-

[11] States also have the option of sharing their state-developed item banks without having them labeled as V-TECS item banks. In this case, V-TECS will help make the banks available to other states, clarifying that these materials may not have been developed according to the standard V-TECS process.

[12] Multiple-choice questions are typically constructed to include three distractor responses, which are plausible alternatives to the correct answer. There is only one correct response. Test takers' responses to these distractors are analyzed in the course of test construction to check on the quality of the questions and answers provided.

nate between examinees who have received instruction in the area being tested and examinees who have received no such instruction.

- *Editing and Completing Final Item Bank.* The item bank is revised based on the field test results.

This process of ongoing, multistage review by workers, V-TECS staff, and other experts helps to ensure the content validity of the items—that they are tied directly to real occupational tasks and duties.

As mentioned above, the banks include both written and performance items. The written tests measure knowledge and skills, with the emphasis on knowledge. There are two categories of written items: those that measure the examinee's ability to recall information, and those that test the examinees's ability to analyze information. V-TECS supplies the correct answer for the written items.

The performance items also measure both knowledge and skills but emphasize skills. Some items focus on the process the student uses, others focus on the product that results from the performance, and some measure both process and product.

Each performance item comes with a description of the task that the student must perform. An evaluator observes student performance using the checklist, checking off actions that the examinee completes successfully.

V-TECS does not supply explicit guidelines on what a successful completion of a task or piece of a task should be like. To understand what is expected of an examinee, the evaluators must refer to the "performance objectives and standards" for each task; which are contained in the V-TECS Analytical Tools. Agencies administering the test can decide on a minimum level of performance required in order to "pass." Scoring is thus performed by the user rather than a testing organization, as in the case of Work Keys.

Use and Availability of Test Item Banks

V-TECS test item banks are used by states for a variety of purposes, including evaluating schools and programs and certifying program or course completion for individual students. Private companies also use some of the V-TECS materials, but not as extensively as schools. Entities using a V-TECS item bank select a number of items from it for each competency to be assessed, and combine them into a test or assessment for use in classrooms.

V-TECS does not maintain records on the number of states, schools, or school districts using its materials, nor does the organization maintain information on which states are using test item banks to meet Perkins Act evaluation requirements. These types of records would be difficult to maintain because the materials are so easily shared.

The V-TECS materials are widely available. Members receive two copies of all V-TECS materials. Nonmember states can purchase the material at a higher cost. In fact, V-TECS materials, including the test item banks, can be purchased by any entity. The only form of marketing V-TECS does is through exhibiting their materials at national conferences. For example, they exhibit at the annual meetings of the American Vocational Association and at the American Society for Training and Development.

The V-TECS organization provides technical assistance to members and other users of its materials. Based on requests from people in member states, V-TECS conducts workshops on such topics as how to write and review items and how to interpret student performances and products. The organization also offers workshops on competency-based vocational education at the request of a state agency or a group of schools.

In addition, V-TECS sponsors periodic national conferences on competency-based assessment and performance standards for vocational technical education. At least once a year, V-TECS holds a workshop for its technical coordinators—

the state agency employees from member states who oversee development of V-TECS material in their state, promote V-TECS materials, and work with the state educators and business people.

The Future of V-TECS

V-TECS is considering whether it should offer a testing service that would actually construct tests for states. In addition, the organization is continuing its efforts to "fill in the gaps" in its current titles. For example, in 1991 V-TECS began identifying related academic skills, a process that it plans to continue for all occupations. Another priority is developing test items for all job areas addressed in V-TECS materials.

NATIONAL OCCUPATIONAL COMPETENCY TESTING INSTITUTE[13]

Overview of NOCTI and the SOCAT Testing Program

NOCTI was created in the early 1970s as a national, not-for-profit educational corporation. All U.S. states and territories are members, although voting privileges are reserved for states that have purchased $100 of NOCTI goods and services during the past year. States are represented through a consortium of vocational education officials. These officials are appointed by state departments of education, state-approved teacher education institutions, and other agencies sanctioned by the NOCTI Board of Trustees. Member states play a role in developing and marketing tests.

The original mission of NOCTI was to develop examinations to assess the occupational competencies of vocational education teachers, an area where there was a paucity of assessment instruments. Today NOCTI is the nation's primary provider of vocational teacher competency exams, offering 60 specific Teacher Occupational Competency Testing (TOCT) examinations.

In the late 1970s, several trends converged to convince the NOCTI board that there was a need for competency tests for vocational students. First, many people felt that competency-based testing would help to improve vocational education and demonstrate its value to employers and the public. There was also emerging interest in developing national standards for vocational student certification.

Responding to these trends, NOCTI formed a consortium of states to develop a system of Student Occupational Competency Achievement Testing (SOCAT) examinations. Since 1978, NOCTI, working with its member states, has overseen the development of 71 SOCAT tests. Each test is tied to a specific occupation, including many of the same occupations covered by the teacher tests. Most of the occupations covered fall within one of the traditional vocational education program areas of business, agriculture, home economics, trade and industry, technical and communications, marketing and distribution, or health.

Using the same methodology, NOCTI also develops Industrial Occupational Competency Testing (IOCT) examinations for specific companies. Industries use the IOCT to conduct preemployment testing, identify in-house training needs, design specific training programs, certify skills, promote employees, and carry out other purposes. Other customers of NOCTI are the U.S. Job Corps and the military.

Somewhat over 50 percent of NOCTI's total revenues come from the IOCT tests for industry. Less than 15 percent of total revenues comes from the student tests (SOCATs), because NOCTI tries to keep the cost of SOCAT tests affordable for schools. The remaining 35 percent or so of revenues comes from the teacher tests and special projects.

[13] The information in this case study is based on OTA interviews with Scott Whitener, president/chief executive officer, NOCTI, and Evelyn Wells, assessment specialist, NOCTI, January and February 1993.

The SOCAT and TOCT Tests

SOCAT test items are based on core competencies considered to be critical for job readiness. These competencies are derived from analysis of the specific tasks performed in a particular job for which the test is being developed. The content of the SOCAT and the TOCT are basically the same for the same job, except for the level of competency. The student tests measure the skills and knowledge expected of a "job-ready" entry-level worker, whereas the TOCT tests measure the skills and knowledge expected of an "experienced" worker. The tests are designed to be used at both the secondary and postsecondary levels.

The SOCAT tests have two components, a written part and a performance part. Both parts are tied to the knowledge and competencies required of job-ready entry-level workers in a given occupation. NOCTI encourages states to use both the written and performance components, because it feels that the written test alone gives an incomplete picture of a student's skills and knowledge. The written tests are multiple choice and are designed to assess whether a student understands and can apply the information needed to perform various tasks. The performance tests require a student to perform various tasks that an individual must be able to do to be considered job-ready. The instructions for doing the performance tests call for having evaluators/administrators from industry judge the quality of both the performance process and the resulting product.

NOCTI also supplies a third test of thinking ability, if requested by the schools. This Test of Cognitive Skills was developed by CTB/McGraw-Hill and is marketed by them. The test measures capabilities of analyzing patterns or sequences of letters, figures, and numbers; recognizing relationships and classifying objects and concepts according to common attributes; recalling previously presented information; and reasoning logically. The items on this test are not tied directly to the SOCAT tests, but are scored by NOCTI and reported back to users as part of the SOCAT score report.

A major difference from V-TECS is that the SOCAT tests are fixed and secure. V-TECS is an item bank. The SOCATs are fixed in that the items are preselected by NOCTI and are the same for all test takers. Clients return completed tests to NOCTI for scoring and do not receive any information about the correct answers. In this respect, the SOCATs are similar to standardized academic tests.

Test Development Process

The process for developing the TOCT and SOCAT tests begins with the analysis of occupational tasks. For each job or occupation addressed, NOCTI assembles a test development team. The team consists of secondary and postsecondary vocational educators from at least two states and industry representatives who are highly competent workers familiar with the tasks required in a given occupation. NOCTI selects some of the industry representatives; state educational agencies select the remaining industry representatives and all of the educators.

These teams then proceed to identify the specific tasks involved in performing the target job or occupation in question by following the DACUM (Developing a Curriculum) process or reviewing existing task lists and merging them into one. Generally, the DACUM method is used only when high-quality task lists do not already exist for an occupation. After prioritizing the tasks, the team members determine which items are best evaluated through a performance exam and which are best done through a written exam. The actual test items are then written either by the team members or subject matter specialists hired for the task.

Once written, the items are reviewed by NOCTI's in-house testing experts. A bias review is conducted and reading experts ensure that the reading level of the exam matches the level required for the occupation in question. This is

accomplished by comparing the test items with actual written materials that workers use on the job.

The draft exam is then pilot- and field-tested in a two-stage process. Teachers or schools first volunteer to administer the test in a small number of schools; NOCTI monitors the administration to make sure the test is operationally sound and that time limits and other administrative features are appropriate. Then, the exam is field tested on a larger scale in at least two states. Through this testing, NOCTI seeks to learn whether the items are free of colloquialisms or regional biases, whether the test items actually match the skills required for the job, whether the time limits are appropriate, and whether the tasks are up to date. This process helps to ensure the validity of the test.

NOCTI also conducts an item analysis to determine which items are not good discriminators among the students—that is, are too easy or too hard for them to answer. Suspect items are reviewed by a subject matter or test expert and possibly revised. Once final refinements and revisions are made, the test is ready to be sold nationally.

SOCAT tests are reviewed annually by NOCTI. After each annual administration, NOCTI asks for feedback from industry representatives, teachers, and students involved in the testing. The feedback is catalogued, and if there is significant criticism of a test, NOCTI reinstitutes a committee of industry representatives and educators to review the test and determine whether revision is necessary.

The technical quality of the test is addressed at several points in the development process. Content validity is achieved by having real workers develop task lists and provide feedback on test content after they are administered. (NOCTI does not check for predictive validity, although the organization would like to do so if more resources were available.)

NOCTI calculates reliability statistics for the written tests using the Kuder-Richardson method for determining the internal consistency of test items. Furthermore, as explained in greater detail below, NOCTI has developed scoring guidelines for the evaluators/administrators of the performance tests.

Use of SOCAT Tests

SOCAT test services may be purchased by any educational agency, including state agencies or individual schools. In 1992, the SOCAT test was administered to 9,015 secondary and postsecondary students in 24 states, with the number per state ranging from 1 in Delaware to 3,435 in Pennsylvania. It is estimated that about one-half of the states using SOCAT tests administer all three components: the written, performance, and cognitive skills tests.

This number of secondary and postsecondary students taking the SOCAT tests is not large in relation to the size of vocational education. In 1992, there were about 720,000 high school seniors who took four credits or more of vocational education in their high school careers and therefore can be considered to be "vocational students."[14]

[14] This estimate was obtained as follows. The number of high school seniors in the 1991-92 school year was 2.4 million. It is reasonable to assume that the number of seniors was about the same in the 1992-93 school year, since the increase in overall high school enrollments between 1991-92 and 1992-93 was less than 2 percent. It is also reasonable to assume that 30 percent of these 1992 seniors were "vocational students." This is the percentage of 1987 high school seniors who took four credits or more of vocational education courses by the time of graduation, and this has been a fairly stable number since the high school class of 1982. A reasonable estimate of the total number of high school seniors who might have taken vocational tests in the 1992-93 school year is therefore 720,000. U.S. Department of Education, Office of Educational Research and Improvement, National Center for Educational Statistics, *Digest of Education Statistics, 1993* (Washington, DC: U.S. Government Printing Office, 1993), tables 2 and 44; and U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, *Vocational Education in the United States, 1969-1990* (Washington, DC: U.S. Government Printing Office, April 1992), table 13.

In some cases, the written SOCAT test is administered as a pretest, and the written and performance tests are used as post-tests to measure the occupational competency gains of students who exit a program, as specified in the Perkins Act. Although NOCTI has traditionally discouraged the use of the written test alone, in 1992 the organization began making the written test available for pretesting because of accelerated interest in using it to fulfill Perkins requirements. NOCTI believes that when the written test is used as a pretest and both the written and performance tests are used as post-tests, valuable information about competency gains can be generated.

In other states, students who do well on the SOCAT test receive special recognition certificates or credentials from the state. These awards are in addition to the SOCAT certificate, regularly provided by NOCTI for each test taker, which carries a printed report of the student's scores on the back. Both the state-issued certificate and the SOCAT certificate are intended for use in student portfolios for interviewing with prospective employers.

In other cases, successful completion of a job-ready SOCAT test may be used to receive advanced standing in college or university programs.

In order to use the SOCAT tests, teachers must ensure that they are teaching the skills and knowledge that the tests measure. Since SOCAT tests are secure, NOCTI regularly provides schools with support materials and test specimen sets to help facilitate test selection.[15]

Sometimes NOCTI helps states to determine whether there is a fit between the SOCAT tests and their state vocational curriculum. For example, NOCTI is currently comparing how the revised task lists developed by a specific state match the SOCAT tasks. In most cases, however, states are responsible for determining whether the tests are compatible with their vocational programs.

Administration of the SOCATs

Member states that use SOCAT tests receive test booklets, test manuals, and videotapes demonstrating the procedures for administering the written and performance examinations. Users can also contact NOCTI for additional technical assistance; NOCTI has a test center in every state with a knowledgeable staff person who can answer questions about the SOCAT tests. In the words of the NOCTI president: "When a school or a state administers the SOCAT tests, they have the whole organization behind them."

The SOCAT tests are designed to be administered by local personnel at local school sites. Any teacher, guidance counselor, or test administrator can administer the written multiple-choice test and the cognitive skills test. However, as mentioned above, NOCTI recommends that a "... journeyworker, trades person, or business representative with technical expertise in the occupation should administer the performance test."

After test administration, a school sends the test results back to NOCTI for machine-scoring. Within 2 weeks of receipt, NOCTI ships two types of reports to the user, individual student reports and teacher reports.

Each test taker receives a student report printed on the back of a certificate, which includes separate scores for the cognitive skills, written, and performance tests. Scores from the cognitive skills and written tests are presented as the percentage of questions answered correctly. Scores from the performance test are presented as the percentage of possible points earned. The student receives an overall percentage score and subscores for different duty areas or tasks. For example, the Electronic Technology written test contains subscores for "schematic symbols," "safety," "soldering," "components," "power

[15] In those situations where the SOCAT tests do not match what is being taught, the NOCTI president contends that the problem is with the curriculum, because SOCAT tests are built around the skills, knowledge, and competencies identified by industry as being necessary.

supplies," and other subjects. The student's written test scores are also compared, by percent, to the scores of other students in the same class, school, and state and with the whole nation.

For the performance tests, students receive ratings for both their performance on individual tasks and the product(s) they produce. The ratings span a scale of 1 through 5, with 1 and 2 unsatisfactory, and 3 through 5 satisfactory. The performance task report lists all of the tasks performed with a single number rating next to each task.

NOCTI has recently developed explicit guidelines and criteria that examiners can use to assign process and product ratings to the test takers' performances. For example, if the student is asked to clean a car assembly, NOCTI now supplies guidelines on how clean the rotor must be for the student to receive a certain score. These guidelines, which are intended to provide inter-rater reliability, are currently available for about one-half of the performance tests. For the remainder of the tests, the evaluator must use subjective judgment to rate student performance.

Teachers receive a class report consisting of a composite of the student reports analyzed by class, school, state, and the nation, along with standard deviations and standard error rates.

Future Priorities

NOCTI is currently developing SOCAT tests for additional occupational areas and plans to continue such expansion. One of the biggest challenges is keeping the existing tests up to date in view of the tremendous changes occurring in industry and in occupational skill requirements.

FINDINGS

The testing products of these three organizations are distinctly different from each other and represent the range of testing practices in vocational education reasonably well. Work Keys is the newest of the three systems and is at one end of the range. V-TECS and the SOCAT tests fit the mold of competency-based testing for job-specific skills, in that both written tests and performance exercises are included, and the skills assessed are job specific. Work Keys is focused on generic workplace skills and is similar to a conventional standardized academic test in the methods of test development that are being employed, the strict procedures of test administration that must be followed, and the closed-ended nature of most of the test items. The SOCAT tests fall in the middle of the range. Like Work Keys, each test consists of a fixed set of items and must be administered according to standardized procedures in order for the resulting test scores to be comparable among test takers. Unlike Work Keys, each test item is directly related to specific work tasks in a particular job area. Work Keys tests employ multiple items to measure one competence that is generally related to all job areas.

Both V-TECS and NOCTI utilize structured methods of job analysis to develop their competency-based materials. Work Keys employs a separate system of job profiling to ascertain the relevance of each general competence to a particular job or group of jobs. V-TECS is different from NOCTI's SOCATs and Work Keys in that it is more a system of resources for competency testing and assessment than a test. V-TECS provides state members of the consortium and local programs within those states with a number of so called "test item banks," from which they construct their own tests reflecting their own local priorities, rather than providing them with tests that are secure and consist of a fixed set of items. Each test item bank is specific to a job area or occupation. Both the NOCTI and V-TECS products include both short answer, written test items, and some performance exercises; however, in both cases the written tests and the performance items are packaged separately and the written components generally predominate.

V-TECS, therefore, does not sell tests but rather models good practices of competency testing and makes testing resources available to

its 23 member states and local programs within those states for them to follow and use in constructing their own tests.

No firm conclusions can be drawn about the validity and reliability of decisions made at the state and local levels using the tests or testing resources produced by these three vendor organizations simply on the basis of the kinds of tests they produce or the methods of test development they follow. Work Keys will have the most data of the three vendors from pilot tests showing the reliability of their instruments. On the other hand, research reviews have found that competency tests of the kind produced by V-TECS and NOCTI are significantly correlated with scores on work sample and hands-on performance tests.[16] It is important to point out, however, that the skills measured by Work Keys are intended to be general across jobs, and the skills measured by V-TECS and the SOCATs are specific.

The extent to which the testing resources produced by these three vendor organizations are currently being used in vocational education differ substantially. The most concrete estimates are for the SOCAT tests, since they must be returned to NOCTI for scoring. The numbers of SOCAT test takers are not large. In fact, NOCTI reports that SOCATs were taken by 9,015 secondary and postsecondary students in 1992. In comparison, the number of high school seniors who were vocational students was about 720,000 in 1992.

The Work Keys system is much newer and no firm estimates of the number of test takers are available yet. At least two states have adopted parts of the Work Keys system, and more are considering the possibilities. Because Work Keys is so different from the SOCAT and V-TECS products in the skills tested and the methods of testing, the effects of Work Keys on testing in vocational education could be substantial, if a significant number of states decide to adopt it. The requirements of the 1990 amendments for performance standards are an important source of states' interest in purchasing Work Keys, as indicated in the interviews conducted by OTA in producing this chapter.

In the state survey conducted by OTA, described in chapter 3, state personnel frequently reported that substantial efforts are devoted to adapting, redeveloping, and/or expanding the competency lists and testing resources produced by V-TECS. The V-TECS materials are used in various ways in these efforts, along with competency lists and competency tests (or test items) from many other sources. The reason commonly given as to why these efforts to adapt and revise materials obtained from elsewhere are necessary is that neither the V-TECS materials nor the materials available from other sources adequately reflect the priorities among different areas of knowledge and skills that are most important in the state or local program area.

How much genuine need exists for this reinvention and adaptation of materials developed elsewhere and how much of it is unnecessary duplication of effort is impossible to say from the data available to OTA. Local priorities among different areas of knowledge and skill undoubtedly differ from state and national priorities, and processes of reinvention have frequently been found to be essential for the thorough implementation of innovations. "To understand is to invent" is perhaps the clearest way of expressing this frequent finding in studies of implementation.[17] On the other hand, questions can be raised

---

[16] Alexandra K. Wigdor and Bert F. Green, Jr. (eds.), *Performance Assessment for the Workplace*, vol. I (Washington, DC: National Academy Press, 1991), ch. 8, and J.E. Hunter, "Causal Analysis, Cognitive Ability, Job Knowledge, Job Performance, and Supervisor Ratings," S. Lundy et al. (eds.), *Performance Measure and Theory* (Hillsdale, NJ: Lawrence Erlbaum, 1983); R.R. Reilly, "The Validity and Fairness of Alternatives to Cognitive Tests," *Policy Issues in Employment Testing* (Boston, MA: Kluwer Academic Publishers, 1993).

[17] Paul Berman and Milbrey Wallin McLaughlin, *Federal Programs Supporting Educational Change, Vol. IV: The Findings in Review*, prepared for the U.S. Office of Education, Department of Health, Education, and Welfare, R-1589/4-HEW (Santa Monica, CA: Rand Corp., April 1975).

about the consequences of this process of reinvention for the comparability of assessment results from place to place and just how necessary and useful it is.

The main conclusion, though, is that the influence of the products of these three vendors of vocational tests on testing and assessment practices in vocational education is limited, at least in relation to all states and all students enrolled in vocational education. V-TECS appears to have the greatest influence through its deliberate strategy of modeling good competency testing and assessment practices for states and local programs to follow, and providing them with competency lists and test item banks to be used as resources in developing their own programs of testing and assessment. However, only 23 states are members of the V-TECS consortium and test item banks are available for only 35 of the over 200 occupational areas in which competency lists are available. V-TECS sells its testing materials to any state or anyone who wishes to buy them. While NOCTI as an organization has many other clients and customers for its testing products, the number of students currently taking their SOCAT test is very limited. Work Keys is too new to know how extensive its impact will actually be, but at least two states (Ohio and Tennessee) have adopted portions of it for statewide use and many more are considering its use.

It is also important to point out that some individual states, such as Oklahoma, which has an extensive program of test development and distribution, also provide competency tests and resources for testing to other states in various ways. The three vendors described here are the most visible vendors of testing resources in vocational education but not necessarily the only such source.

Broad Technical Skills 5

The world of work is changing rapidly, and in ways that are difficult to anticipate. In response to this change and in an effort to make vocational education a method of equipping young people for more than one occupational future, policymakers and educators have begun to seek skill training and preparation that applies to many jobs. The language of the 1990 amendments to the Carl D. Perkins Vocational Education Act refers to "broad technical skills." The Office of Technology Assessment (OTA) was asked to determine how much testing and assessment was aimed at measuring the mastery of these skills. This concept is a new one, however, and no agreement exists on what skills are truly lifetime occupational skills, or on how those skills are best acquired.

Very few states are moving toward organizing their vocational programs around broad technical skills. One of the principal reasons for this is the lack of research and proven experience showing what broad technical skills are, how they can be taught and assessed, and what the advantages and disadvantages of such a change would be.

The implications of organizing vocational programs around broad technical skills are potentially far reaching. Large segments of the training enterprise in this country are oriented to providing people with just the skills needed for particular jobs instead of careers. By organizing more of training around the development of broad technical skills, a greater proportion of the total training effort might be directed to the preparation of people for careers and learning on the job rather than for specific entry-level positions. This lengthening of the "skill horizon" of vocational education and other training efforts would require

placing greater emphasis on developing capabilities for thinking, responding to the unexpected, and taking initiative, in addition to acquiring the skills needed for performing specific occupational tasks.

Broadening the technical and occupational skills around which vocational education is organized could also help to facilitate the integration of academic and vocational education. By shifting the content of vocational preparation toward the development of capabilities that require more thinking and knowledge application, opportunities should expand for connecting occupational and academic content.

Whether the potential advantages from broadening the technical skills concept can be realized is an open question. Certain major fields of vocational education have historically been much more broadly oriented than others. Agriculture is a clear example, while the trades have been more specialized. The recent pressures on all fields, including agriculture, have been to become more competency based and job oriented.

The main purpose of this chapter is to suggest some alternative ways of beginning to define broad technical skills. Much more thought needs to be given to this basic concept than has been given to it so far.

Broad technical skills are different from generic workplace skills in certain key respects. The main difference is that generic workplace skills have been defined as essentially the same across industries. Whether people who have these generic workplace skills in one industry or occupational area in fact demonstrate them with equal proficiency in other occupational or industry areas has not been shown.

The concept of broad technical skills reflects much more directly the content of technology, information, and methods of organization within an industry or group of industries. It is possible that these industry groups could be as broad as health care, financial services, manufacturing, hospitality, and agribusiness; or it may be necessary to limit them to occupational areas within industry groups. Both broad technical skills and generic workplace skills may be important for productivity in the workplace and career development. Broad technical skills will be defined in this report to include not only the specific technical content of job performance but background knowledge of the technological, social, and economic development of the industry.

Five different approaches to defining broad technical skills are described in this chapter. The approaches are: vocational abilities and aptitudes, occupational maps, core occupational skills, design and technology, and cognitive problem solving. They are founded on fundamentally different assumptions about the nature of skills, relationships between skills and the contexts in which they are used, and relationships between general and specific skills. The approaches differ greatly in the extent to which they are supported by research; some reflect extensive analysis and others exist primarily as practices or even ideas for improvements in practice. These approaches do not exhaust the range of possibilities.

VOCATIONAL APTITUDES

One approach to defining broad technical skills is to think of them as vocational aptitudes. The thesis of vocational aptitude testing is that people perform best in the jobs or occupational fields for which they have the most "abilities" and that these abilities can be identified through tests. The instruments developed for assessing vocational aptitudes employ the techniques of testing for achievement and multiple abilities developed from research in psychology on mental testing and intelligence.[1] Aptitude tests differ from tests of achievement mainly in the uses that are made of the results rather than the abilities measured

[1] Lee J. Cronbach, *Essentials of Psychological Testing* (New York, NY: Harper and Row, 1990), ch. 10.

and the nature or content of the tests employed.[2] Aptitude tests are constructed to measure the likelihood that a person will perform well in the future in different kinds of jobs or occupational fields, while achievement tests are primarily intended to measure what a person has learned in response to education, such as a training program for computer technicians or mathematics up to a certain grade in high school.[3]

Most aptitude tests consist of several individual tests for different areas of knowledge or ability. The tests are selected because of their contribution to the accuracy with which performance in a subsequent job or occupation can be predicted. These component tests of the overall test "battery" typically include measures of what can be considered pure cognitive ability, such as memory span, spatial reasoning ability, or perceptual speed; academically related achievement, such as vocabulary, reading comprehension, or arithmetic problem solving; perceptual-motor skills, such as manual dexterity or eye-hand coordination; and other areas of knowledge or skill more specific to the occupations of interest.[4] These four domains of ability are different from a person's vocational interests and personality characteristics, which can also be measured through another kind of test called an "interest inventory," such as mechanical comprehension or computer knowledge. Both vocational aptitude tests and interest inventories have been shown to be significantly related to job performance.[5]

Tests of vocational aptitude are developed primarily for purposes of career counseling or screening people for selection into jobs or occupational areas. Because of their origins in mental testing, the format of most aptitude tests is written, except for measures of perceptual-motor skills. Because of the written format, aptitude tests can be group administered and are therefore relatively inexpensive. This is essential given the primary use of this kind of test, which is screening large numbers of job applicants.

One view is that the knowledge and abilities measured through vocational aptitude testing are prerequisites for success in different kinds of jobs. Another interpretation is that measurement of knowledge and abilities provides an indication of the individual's capability for learning new tasks that are similar to those that will be performed on the job. This latter interpretation opens a door to making connections among the different approaches to broad technical skills.

Vocational aptitudes differ from job competencies in that they are presumed to underlie people's performance of tasks on the job. Job tasks describe specific behaviors that people exhibit in response to certain work goals or

---

[2] "Achievement and aptitude tests are not fundamentally different. . . . Tests at one end of the aptitude-achievement continuum can be distinguished from tests at the other end primarily in terms of purpose. For example, a test for mechanical aptitude [could] be included in a battery of tests for selecting among applicants for pilot training since knowledge of mechanical principles has been found to be related to success in flying. A similar test [could] be given at the end of a course in mechanics as an achievement test intended to measure what was learned in the course." National Academy of Sciences Committee on Ability, *Ability Testing: Uses, Consequences, and Controversies: Part I: Report of the Committee* (Washington, DC: National Academy Press, 1982), p. 27.

[3] Anne Anastasi, *Psychological Testing*, 6th ed. (New York, NY: Macmillan Publishing Co., 1988), p. 412.

[4] Psychologists frequently distinguish between "fluid" intelligence, such as spatial ability, visualization ability, short-term memory, and so forth; and "crystallized" intelligence, which is the product of experience, such as arithmetic reasoning and word knowledge. The theory is that people use their fluid abilities to obtain crystallized abilities. See, for example, Richard E. Snow, "Aptitude Theory: Yesterday, Today, and Tomorrow, *Educational Psychologist*, vol. 27, No. 1, 1992, pp. 5-32. A wide range of predictor variables for vocational aptitude testing is discussed in Norman G. Peterson et al., "Project A: Specification of the Predictor Domain and Development of New Selection/Classification Tests," *Personnel Psychology* vol. 43, No. 2, summer 1990, pp. 247-276.

[5] Generally, the findings are that cognitive and perceptual-psychomotor ability tests provide the best prediction of job-specific and general task proficiency, while measures of temperament and personality are the best predictors of giving extra effort, supporting peers, and exhibiting personal discipline. Scores derived from interest inventories correlate more highly with task proficiency than with demonstrating effort and peer support. Jeffrey J. McHenry et al., "Project A Validity Results: The Relationship Between Predictor and Criterion Domains," *Personnel Psychology*, vol. 43, No. 2, summer 1990, pp. 335-354.

demands of the job environment, and job competencies are tasks that people have demonstrated that they are able to perform repeatedly. The concept of vocational aptitudes is that the performance of these specific tasks requires a certain combination of abilities, or traits, as they are sometimes called, and that different tasks require different combinations of these abilities or traits. Their mental basis is what lends these traits or abilities their presumed generality of application across many different kinds of specific tasks. Job competencies, on the other hand, are not generally presumed to be transferable across jobs or work situations in the ways that they are usually defined. There is no model of transfer built into the job competency model. Aptitudes are thus at one extreme essentially fixed (though learnable) qualities of the mind, while job competencies (or job tasks) describe highly specific, observable job behaviors that vary greatly from job to job. A "giant leap" of faith is required to translate between the two.[6]

Two approaches to filling this gap are described later in this chapter. One is occupational mapping, which involves broadening the definition of job tasks to cover whole clusters of occupations and revising the nature of the task descriptions to include thinking as well as behavior. The other approach is the cognitive skills model, which provides a detailed description of the thinking processes involved in acquiring the knowledge required in a job domain, and becoming expert at performing a range of tasks in the domain and responding to a variety of new task demands.

Vocational aptitudes are important to consider in defining broad technical skills because of the way the tests are developed. A number of different areas of knowledge and abilities are hypothesized to be significant predictors of performance in a range of jobs or occupational areas. Each of the abilities is then described in terms of a "construct" or hypothesized relationship between processes of thinking and observable aspects of performance. Subtests for each of these constructs are then developed and assembled into a larger test for administration to a sample of individuals to obtain information about how their scores on the subtests are related to measures of their performance. Methods of statistical inference are employed to determine which of the subtests provide the greatest capability for distinguishing between high and low performers across the range of jobs or occupations selected. Constructs found through this process to be highly correlated with performance that is "technical" in nature can then be considered candidates for broad technical skills.

One important example of a vocational aptitude test is the Armed Services Vocational Aptitude Battery (ASVAB). The ASVAB is used by the military for selecting personnel into the armed forces and assigning them to jobs and initial training. It consists of 10 subtests, as shown in table 5-1.

The military also provides the ASVAB free of charge to high schools as the main part of a nationwide student career counseling program. Over a million high school students take the ASVAB every year (out of, for example, the approximately 2.4 million seniors in 1992[7]). The ASVAB is administered in the schools by personnel from the Military Entrance Processing Command and scored by them. Each school receives in return a report on students' performance and each student a complete profile of his or her scores, a guide to military careers, and carefully

[6] Alexandra K. Wigdor and Bert F. Green, Jr. (eds.) *Performance Assessment in the Workplace, Vol. I* (Washington, DC: National Academy Press, 1991), p. 85.

[7] See footnote 14 in ch. 4 of this report.

Table 5-1—Subtest Content Areas of the Armed Services Vocational Aptitude Battery

Subtest content area	Description of the subtest
General science (GS)	Knowledge of the physical and biological sciences
Arithmetic reasoning (AR)	Word problems emphasizing mathematical reasoning rather than mathematical knowledge
Word knowledge (WK)	Understanding the meaning of words, i.e., vocabulary
Paragraph comprehension (PC)	Presentation of short paragraphs followed by multiple-choice items
Numerical operations (NO)	Speeded test of four arithmetic operations
Coding speed (CS)	Speeded tests of matching words and four-digit numbers
Auto and shop information (AS)	Knowledge of auto mechanics, shop practices, and tool functions in verbal and pictorial terms
Mathematics knowledge (MK)	Knowledge of algebra, geometry, and fractions
Mechanical comprehension (MC)	Understanding of mechanical principles, such as gears, levers, and pulleys
Electronics information (EI)	Knowledge of electronics and radio principles in verbal and pictorial terms

SOURCE: John R. Welsh and Susan K. Kucinkas, *Armed Services Vocational Battery: Integrative Review of Validity Studies*, AFHRL-TR-90-22 (Brooks Air Force Base, TX: Human Resources Laboratory, July 1990), table 2.

prepared instructions explaining how they can use the information to decide which military occupations they are best qualified to pursue.[8]

Scores on the ASVAB have been shown to predict job performance reasonably well using a number of different criteria: grades in training school; scores on job-specific, hands-on performance assessments; scores on job knowledge tests;[9] supervisor ratings; field proficiency ratings; and many other variables.[10] The multiple correlation between ASVAB scores and normalized scores on the job performance assessments is typically in the neighborhood of 0.6, which the National Academy of Sciences has concluded is a "useful amount."[11] This correlation of 0.6, with some variation higher and lower, has held across many different kinds of military jobs[12] and to some extent equivalent civilian jobs.[13] The total amount of variance in the criterion variable explained by this correlation of 0.6, whether it is grades in training school or hands-on job performance, is

[8] U.S. Department of Defense, Defense Manpower Data Center, *Counselors Manual for the Armed Services Vocational Aptitude Battery*, DOD 1304.12-L-ASTP-CM (North Chicago, IL: U.S. Military Entrance Processing Command, July 1992).

[9] Job knowledge tests are essentially the same type of test as the job competency test discussed in ch. 3.

[10] Charlotte Campbell et al., "Development of Multiple Job Performance Measures in a Representative Sample of Jobs," *Personnel Psychology*, vol. 43, No. 2, summer 1990, pp. 277-300. See also Paul W. Mayberry and Neil B. Carey, *Relationship Between ASVAB and Mechanical Maintenance Job Performance*, CRM 92-192 9 (Alexandria, VA: Center for Naval Analysis, March 1993). An extensive set of hands-on job performance measures was developed for the joint service Job Performance Measurement project overseen by the National Academy of Sciences, which are utilized in these and many of the other studies reported in this chapter.

[11] Wigdor and Green (eds.), op. cit., footnote 6, p. 163.

[12] Ibid., p. 151; Milton Maier, *Military Aptitude Testing: The Past Fifty Years*, DMDC Technical Report 93-007 (Monterey, CA: Defense Manpower Data Center, June 1993), p. 6; and Paul W. Mayberry, *Validation of ASVAB Against Infantry Job Performance*, CRM 90-182 (Alexandria, VA: Center for Naval Analysis, December 1990). The predictive validity ranges from 0.2 or higher to over 0.70, but in most jobs is about 0.6. Ibid., p. 1263. These correlations are all corrected for "restriction of range," according to the procedures recommended in Stephan Dunbar and Robert Linn, "Range Restriction Adjustments in the Prediction of Military Job Performance," *Performance Assessment in the Workplace, Vol. II*, Alexandra K. Wigdor and Bert F. Green, Jr. (eds.) (Washington, DC: National Academy Press, 1991), pp. 127-157. The correction for restriction of range involves adjusting the computed correlation between ASVAB scores and the criterion variable for the fact that performance scores are available only for individuals who have been selected for the military. Without correcting for the restriction of range, the predictive validity of the ASVAB is in the neighborhood of 0.4. Corrections are also sometimes made for unreliability of the criterion measure in some analyses of predictive validity.

[13] R.L. Holmgren and M.R. Dalldorf, *A Validation of the ASVAB Against Supervisors' Ratings in Civilian Occupations* (Palo Alto, CA: American Institutes for Research, 1993).

generally about 25 percent.[14] The predictive validities of the ASVAB are somewhat higher for grades in training school than for hands-on performance assessments (0.1 to 0.2 higher), and somewhat lower but much more variable for field proficiency ratings and supervisor ratings.[15]

Research in the Marines shows that the predictive value of the ASVAB continues after the first term of enlistment for both infantry and leadership personnel. The main finding is that over 3 years of experience are required for personnel with low tested aptitude on the ASVAB to reach the same level of proficiency on the job performance assessments as high aptitude personnel demonstrate in their first year.[16]

The critical question for the purposes of this chapter is the contribution of the three technically oriented aptitudes to job performance in comparison to the academically and cognitively oriented aptitudes, and the extent to which these technical aptitudes are more critical for performance in some occupations than others.

Factor analysis shows that scores on the 10 different tests of the ASVAB can be reduced to four intercorrelated factors: verbal aptitude (general science, word knowledge, and paragraph comprehension), mathematical aptitude (arithmetic reasoning and mathematics knowledge), cognitive speed (numerical operations and coding speed), and technical aptitude (auto and shop information, mechanical comprehension, and electronics information).[17] The issue for this report is how much ability in this fourth domain of technical knowledge and skills contributes to the predictive validity of the ASVAB beyond what is contributed by the more academically and cognitively oriented domains.[18]

The actual content of the tests for technically oriented aptitudes is illustrated in box 5-A. The test questions shown indicate the nature of the knowledge and skills tested in these areas.

Hunter et al. have asserted that the 10 aptitude domains of the ASVAB can be reduced to only 1 factor of general ability, because they do not contribute separately to the predictive power of the ASVAB across different military occupations.[19] Their research shows that statistical differences in the numbers of individuals sampled and other such "artifacts" of the research designs account for all of the differences previously found in the ability of subtests of the ASVAB to predict performance among occupations. By correcting for all of these statistical artifacts, no more variance in performance scores could be explained using separate scores on the 10 different subtests than could can be explained by a single

---

[14] D. R. Divgi and Paul W. Mayberry, *The Role of Aptitude Factors in Predicting Hands-on Job Performance*, CRM 69-215 (Alexandria, VA: Center for Naval Analysis, March 1990), p. 5; and Lauress Wise, "The Validity of Test Scores for Selecting and Classifying Enlisted Recruits," *Test Policy in Defense: Lessons From the Military for Education, Training, and Employment*, Bernard Gifford and Linda C. Wing (eds.), National Commission on Testing and Public Policy (Boston, MA: Kluwer Academic Publishers, 1994), table 15, pp. 221-260.

[15] John E. Hunter, "A Causal Analysis of Cognitive Ability, Job Knowledge, Job Performance, and Superior Ratings," *Performance Measurement and Theory*, F. Landy et al. (eds.) (Hillsdale, NJ: Erlbaum, 1993). See also Wigdor and Green (eds.), op. cit., footnote 6, table 8-10; and Mayberry, op. cit., footnote 12, table 3.

[16] Paul W. Mayberry, *Analysis and Prediction of Infantry Unit Leaders' Performance*, CRM 91-99 (Alexandria, VA: Center for Naval Analysis, June 1991), figure 1; and Mayberry, op. cit., footnote 12, figure 1.

[17] Peter H. Stoloff, *Factor Analysis of ASVAB Form 8a in the 1980 DOD Reference Population*, CNA Memorandum 83-3135 (Alexandria, VA: Center for Naval Analysis, August 1983).

[18] The ASVAB currently has no test for perceptual-psychomotor or spatial abilities, as do other aptitude tests like the General Aptitude Test Battery. Consideration is being given to adding "assembling objects" as a spatial reasoning test to the ASVAB.

[19] John E. Hunter et al., "The Validity of the Armed Services Vocational Aptitude Battery (ASVAB) for Civilian and Military Occupations," U.S. Air Force Contract No. F416689-83-C-0025, August 1985. See also Lee J. Cronbach, "Five Perspectives on Validity Argument," *Test Validity*, Howard Wainer and Henry J. Braun (eds.) (Hillsdale, NJ: Lawrence Erlbaum Associates, 1988), pp. 3-18; Cronbach, op. cit., footnote 1, pp. 293-398; and Malcolm Jones and James A. Eacles, "Predicting Training Success: Not Much More Than g," *Personnel Psychology*, vol. 44, 1991, pp. 321-332.

Box 5-A—Examples of the Content of Test Items on Subtest of the ASVAB for Broad Technical Skills

Auto and Shop Information—

- A car uses too much oil when which parts are worn?
  (a) pistons (b) piston rings (c) main bearings (d) connecting rods
- The saw shown above is used mainly to cut:
  (a) plywood (b) odd-shaped holes in wood (c) along the grain of the wood
  (d) across the grain of the wood

Mechanical Comprehension—

- In this arrangement of pulleys (a system of pulleys with different diameters is shown), which pulley turns faster?
  (a) pulley A (b) pulley B (c) pulley C (d) pulley D
- Which post holds up the greater load? (a load at one end of a beam supported by two posts is shown)
  (a) post A (b) post B (c) both equal (d) not clear

Electronics Information—

- Which of the following has the least resistance?
  (a) wood (b) iron (c) rubber (d) silver
- In the circuit shown (a battery connected to a resistor) the resistance is 100 ohms and the current is 0.1 amperes. What is the voltage?
  (a) 5 (b) 10 (c) 100 (d) 1,000 volts

SOURCE: U.S. Department of Defense, Defense Manpower Data Center, *Counselor Manual for the Armed Services Vocational Aptitude Battery*, DOD 1304.12-L-ASTP-CM (North Chicago, IL: U.S. Military Entrance Processing Command, July 1992), pp. 131-135.

factor of general ability. This is called the validity generalization hypothesis.

More recent research using new hands-on measures of job performance contradicts this hypothesis with regard to technical aptitudes. This newer research shows that the contribution of technical aptitude to the predictive validity of the overall test battery is significant. This implies that the technical aptitude measured by the ASVAB is different from the more purely cognitive and academically oriented areas of the test. Using new testing methodologies developed for the Marine Corps Job Performance Measurement Project, which was overseen by a committee of the National Academy of Sciences, Divgi and Mayberry have recently found in a study of Marine personnel that the technical aptitude composite of the ASVAB explains 10 percent of the total variation in performance explained by the whole test.[20] The statistical model they used first accounts for all of the variation in performance that can be accounted for by one general factor of ability and then, secondarily, for all of the additional variation that can be accounted for by a composite of the three technical domains considered together (auto and shop information, mechanical comprehension, and electronics information). This is a conservative estimate of the separate effects of the technical aptitude composite on performance. They conclude that one of main reasons why Hunter et al. found that general ability accounts for all of the variance in performance scores is that they used grades in training school as their criterion variable rather than hands-on performance tests. When grades in training school are used as the criterion variable,

[20] Divgi and Mayberry, op. cit., footnote 14, table 2.

Divigi and Mayberry find that the technical composite adds little to the predictive validity of the ASVAB.[21]

Similar results were found in another study of Marine infantry by the National Academy of Sciences. When job knowledge tests were used as the criterion variable, the three major, "non-speeded" factors of the ASVAB have about the same predictive validities of about 0.73 (the three non-speeded composites are mathematical, verbal, and technical aptitude).[22] When hands-on performance tests are used, the predictive validities of the same three factors varied significantly. The correlation of technical aptitude with the hands-on performance test results was the highest among the three composites at 0.65, while the other two were about 0.57.

The validity of the ASVAB for predicting performance in mechanical occupations in the Marines has been studied by Mayberry and Carey. They find that, for hands-on performance tests, the predictive validity of the mechanical maintenance composite of the ASVAB used by the Marines[23] averaged 0.068 higher than the next best composite, which is General Technical (GT),[24] and over 0.15 higher than the Armed Forces Qualification Test (AFQT), which is a measure of general ability used by the military.[25] These translate to an average of 11 percent above GT and 30.6 percent above the AFQT. The other composites were clerical and electronics.[26] No such differences were found in predicting job knowledge test scores or grade point averages in training school. In other words, the technical aptitude measured through performance tests is different from general intelligence, while the technical aptitude measured through written job competency tests or grades in training school is not.

Another source of evidence for the differential validity of the ASVAB can be obtained by comparing the predictive validity of the subtests across occupations, as shown in figure 5-1. The validities of the different subtests are higher in occupations where the nature of the job tasks corresponds with the nature of the subtest. Evidence like this does not prove the differential validity of the ASVAB but strongly suggests that it exists.[27] For infantryman, there are few differences in predicted hands-on performance among the 10 individual subtests, except for mechanical comprehension, which is noticeably higher than the rest. This agrees with the findings of Mayberry and Carey. For administrative occupations, the validities of arithmetic reasoning and mathematical knowledge are substantially higher than for all of the other subtests and the other occupations. For mechanical jobs, the technical composite is by far the highest and all of the other validities are lower than for all other occupations.

These results from psychometric studies have been confirmed using wage data for civilian occupations. Using the National Longitudinal Study, one researcher found that both wages and earnings in civilian occupations are significantly correlated with scores on the mechanical, elec-

---

[21] Ibid., p. 9.

[22] Wigdor and Green (eds.), op. cit., footnote 6, figure 8-4. The study was based on P.W. Mayberry, *Interim Results for Marine Corps Job Performance Measurement Project*, Research Memorandum 88-37 (Alexandria, VA: Center for Naval Analysis, 1988).

[23] The mechanical maintenance composite consists of the total ASVAB scores for arithmetic reasoning, auto and shop information, mechanical comprehension, and electronics information.

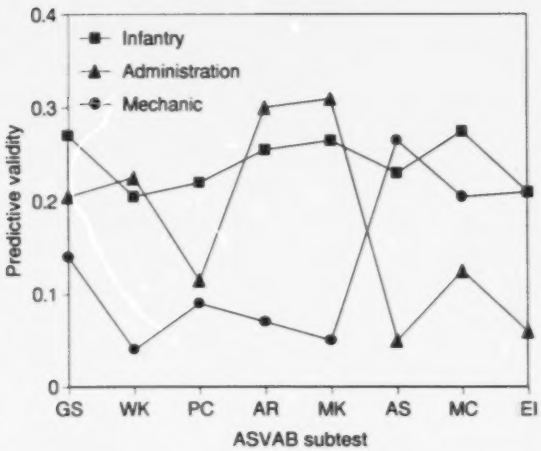
[24] The General Technical composite consists of arithmetic reasoning, word knowledge, paragraph comprehension, and mechanical comprehension.

[25] Mayberry and Carey, op. cit., footnote 10, table 13. The predictive validities of the mechanical composite were 0.64 or greater for four of the five mechanical occupations studied and one validity was 0.70.

[26] Clerical consists of the mathematics knowledge, word knowledge, paragraph comprehension, and code speeding subtests, and electronics is arithmetic reasoning, general science, mathematics knowledge, and electronics information.

[27] Wigdor and Green (eds.), op. cit., footnote 6, p. 171.

Figure 5-1—Correlations of ASVAB Tests With Hands-On Performance Test Scores (Army occupations)



KEY:

GS = General science	MK = Mathematics knowledge
WK = Word knowledge	AS = Auto/shop
PC = Paragraph comprehension	MC = Mechanical comprehension
AR = Arithmetic reasoning	EI = Electronics information

SOURCE: Alexandra K. Wigdor and Bert F. Green, Jr. (eds.), *Performance Assessment for the Workplace, Vol. I* (Washington, DC: National Academy Press, 1991), p. 171.

tronics, and shop information subtests of the ASVAB.[28] No effects were found for scores on the academic and cognitive subtests, but they are highly correlated with years of schooling, which were included in the analysis as a control variable. The independent correlation of the technical aptitude with wages provides strong indication of its independence of the academic and cognitive aptitudes of the ASVAB.

Considered altogether, these studies provide substantial evidence that the technical aptitude measured by the ASVAB, whose content is represented by the test questions in box 5-A, is significantly related to job performance and substantially different from academic and purely cognitive aptitudes.

OCCUPATIONAL MAPS

A second concept of broad occupational skills is being developed by the American Electronics Association (AEA) with support from the U.S. Department of Labor. The effort is one of several pilot projects being supported by the Departments of Labor and Education to develop and demonstrate the feasibility of industry skill standards. In these projects, grantees are urged to organize their standard-setting efforts around what the government has called "broad occupational areas." For the purposes of this report, the approach will be called "occupational mapping."

AEA's approach to skill standards involves reconstructing the job competency model around generically defined job tasks that cut across a range of occupations, and including categories of related knowledge and technical skills. The new model is intended to provide the industry with the means to "speak with one voice" on the skills needed for careers within industry and yet reflect the major differences in specific contexts of employment and jobs. Under the approach, it must be possible to demonstrate the technical accuracy of the resulting skill standards—that the standards are related to high performance in the workplace. The skill standards are also expected to be useful for a wide variety of human resource planning and development needs within the industry other than certification of the competence of individuals.

The orientation of the AEA approach to defining job tasks is borrowed from the job competency model. However, the tasks are identified from the top down within broad clusters of industry jobs rather than the bottom up within narrow categories as in the job competency model. For the AEA project, three broad clusters of occupations have been chosen: manufacturing specialist, pre- and post-sales analyst, and administrative/information services support analyst. AEA estimates that employment in these

[28] John Bishop, Center for Advanced Human Resource Studies, Cornell University, "Educational Reform and Technical Education," working paper No. 93-04, 1993.

Box 5-B—Broad Occupational Tasks for Manufacturing Specialists

Key Purpose

Develop, manufacture, deliver, and improve electronics-related products and processes that meet or exceed customer needs.

Critical Functions

Ensure production process meets business requirements.
Initiate and sustain communication processes and procedures.
Establish customer needs.
Determine design workability and manufacturability.
Use human resources to manage work flow.
Select, obtain, and optimize available machines and equipment to meet product process requirements.
Make products that meet customer specifications.

Includes Jobs Such as

Production associate, operator, production technician, and assembler.

Scenario

Manufacturing specialists are on the front line of the manufacturing operation. They work as individuals, or increasingly, as members of "self-managed" teams. They often play a key role in work scheduling and resource management within their work group.

They use a variety of quality, cost-reduction, and cycle-time productivity tools and processes. Their direct communication with internal and external customers and other specialist staff is increasing.

SOURCE: American Electronics Association, draft internal memorandum, November 1993.

three occupational areas covers about 60 percent of all jobs at skill levels not requiring a baccalaureate degree in the computer, software, semiconductor, consumer electronics, and telecommunications industries.

This "mapping" of the broad occupational clusters into tasks is done by industry representatives. The process involves first identifying the major purpose of work in each area and a limited number of "critical functions" defining the content of work output. An example of the major purpose and the critical functions for manufacturing specialists is shown in box 5-B.

The next step in the mapping process is to identify a limited number of activities that must be performed in order to achieve each of the critical functions in the cluster area and explicit criteria for knowing when the activities are being performed well. These criteria will provide the basis for setting performance-level standards for particular jobs within the three broad industry clusters. The specific standards adopted will differ among job areas and segments of the industry, but will remain consistent within the task-oriented framework. In the end, the mapping process will result in skill definitions similar to the job competency model, but guided by a framework of critical functions and performance criteria. Each critical function, together with the associated set of activities and performance criteria, is called a competency module. These competency modules will be the basic unit of skill certification. An example of the competency module for the second critical function of manufacturing specialists is shown in box 5-C.

The criteria defining when critical functions are performed well are one of the major ways in which the occupational mapping strategy of AEA differs from the job competency approach. In the job competency model, the criteria are frequently

BEST COPY AVAILABLE

Box 5-C—Activities and Performance Criteria for Manufacturing Specialists: Ensure Production Process Meets Business Requirements Competency Module

Integrate Improvement Process Into Each Critical Function
- Quality monitoring and improvement processes are performed and are documented according to company procedures.
- Deviation and root cause of deviation are identified from ongoing analyses of processes.
- Recommendations for process improvement are documented, approved, and implemented.

Meet Health, Safety, and Legal Requirements With Regard to Process, Product, and People
- Health and safety requirements and procedures are implemented and followed at all times.
- Potential health and safety hazards are identified through continuous safety review.
- Confidentiality of proprietary information is protected according to company policy.
- Company standards of business conduct are followed.

Select Setup and Perform Diagnostic Tests
- The selected test method meets product specifications and customer requirements.
- The test method is safe, cost-effective, and meets time needs.
- Equipment setup conforms to required test and space specifications.
- Test equipment is calibrated correctly and functions according to specifications.
- Proper handling procedures are followed.
- Tests and test documentation are completed according to prescribed sequence, time, and quality requirements.
- Test results and serialization of tested products are accurately documented.

Analyze and Interpret Test Data for Problems That Require Corrective Actions and for Compliance With Specifications
- Products or process deviations and root causes of deviations are accurately identified and documented, and corrective action is initiated.
- Systems for evaluating remedial action are established.
- Corrective action and appropriate recommendations are documented.
- Products forwarded to customer on a "conditional accept" basis are accompanied by accurate documentation completed to customer specifications.
- Tests are in compliance with legal requirements, company policy, and customer specifications.

SOURCE: American Electronics Association, draft internal memorandum, November 1993.

only restatements of the tasks,[29] or a set of simple terms such as "skilled," "moderately skilled," or "unskilled," as was discussed in chapter 3. This aspect of the occupational mapping approach provides a clearer methodology for setting performance-level standards than does the job competency method.

As indicated by the manufacturing specialist example, the broad task structures being identified by AEA are also different from the job competency model in the extent to which they reflect the basic concepts of high-performance work systems. The overall structure of the critical functions for manufacturing specialists clearly reflects, for example, the responsibilities being given to front-line workers for monitoring quality, continuous improvement, and dealing with customers.

The next phase of the AEA project will be to identify the categories of technical skills, knowledge, and understanding that underlie performance of the critical functions. In these areas, AEA

[29] For example, if the task is "prepares checklists," the criteria is "prepared checklists."

BEST COPY AVAILABLE

will incorporate methodologies for measuring skills and abilities, as discussed in other sections of this chapter. These could include ways to measure the technical knowledge or aptitudes of individuals, or their cognitive skills. This knowledge and underlying capabilities aspect of occupational mapping indicates how the technique may provide a framework for integrating different approaches to broad occupational skills. It also is where AEA is considering how the skills structure can relate to the K-12 system of education. Similar approaches to the definition of broad occupational skills have been followed in Great Britain and Australia.

CORE OCCUPATIONAL SKILLS

As shown in figure 3-9 in chapter 3, five states are planning to adopt or have adopted some strategy for organizing their programs of vocational education around broad technical skills. In most of these cases, the strategy adopted by the state is to divide vocational education into broad clusters of occupations and, in the early years of preparation, organize students' programs around the development of some set of core occupational skills. Each core area of knowledge and skills then branches at the next step into sequences of clusters of courses that become increasingly more specialized. Some examples of these broad cluster areas are the traditional areas of business, agriculture, and health. States typically have 6 to 10 cluster areas covering all industries.

New York was the first state to adopt such a statewide policy, starting in 1983.[30] Through a long process of planning and development, a "continuum" of occupational education was adopted. A key aspect of the continuum is that, starting in the earliest grades and continuing throughout high school, the occupational curriculum is guided by learning objectives in five broad areas:

1. personal development (e.g., the development of a positive self-concept, learning to work in groups, and learning how to find a job);
2. social systems (e.g., understanding the economics of business and the legal system);
3. information skills (e.g., communicating orally and using computers to present information);
4. resource management (e.g., managing money and time); and
5. technology (e.g., understanding systems in manufacturing, construction, communications, and agriculture).

Occupational education starts in the elementary grades with courses in the use of computers. By the end of the 8th grade, all students must have taken required courses in keyboarding and computer literacy, home and career skills, and Introduction to Technology.[31]

All students who decide to major in vocational education are then required to take a 1-year course in the Introduction to Occupations. This course is organized around the five core occupational competencies, with an emphasis on employability and other generic workplace skills in the first half of the course and more occupationally specific knowledge and skills in seven occupational cluster areas in the second half. In the first half of the Introduction to Occupations course, students take units in the economics of business and personal finances, human relations, decisionmaking and problem solving, energy systems, and other topics. The seven cluster areas that start in the second half are agricultural education, business/marketing education, health occupations education, home economics education, technical educa-

[30] This description of the New York program is based on Richard Jones, former Director of the Division of Occupational Education Instruction, personal communication, Aug. 24, 1993.

[31] The Introduction to Technology course is similar to the curriculum described in the next section on Design and Technology, but the focus is on learning about the systems of technology.

tion, technology education, and trade and industrial education. In agricultural education, for example, there is a module in the second half of the Introduction to Occupations course on plant science that covers the basic biology and agronomy of plants (seeds, roots, leaves, stems, and flowers), life supporting processes and reproduction, environmental factors, soil and fertilizers, plant pests, and so forth. Following the Introduction to Occupations course, students take a subject area core course and sequences of three or more increasingly more specialized courses in one of these seven cluster areas. Instruction at these more advanced levels is still guided by the same core occupational skills but at more advanced levels of learning.

An assessment program has been implemented to sustain the teaching of the core occupational competencies throughout the curriculum. The test for the Introduction to Occupations course consists of a locally administered but state-developed competency test. The state requires that all students who complete the course take the test. The state also has competency tests in four of the seven occupational cluster areas, which are taken on completion of required coursework in those areas.

California is in the process of developing and implementing a similar approach to organizing vocational education around core skills and occupational clusters. For example, in business education, which is serving as a model for the other cluster areas, a framework of curriculum standards will soon be adopted for the first year course of the program. This course is called the business technology core. After completing the business technology core, students will choose one of four "career path clusters" in computer science and information systems, business administration, accounting and finance, or marketing. Three sequential levels of learning outcomes are planned within each cluster; these will be adopted as a series of standards. There will be career path cluster standards, career path specialization standards, and an entrepreneurship standard. The business technology core will also be linked to instructional units or courses in business exploration planned in the upper elementary grades.

Cutting across all of these cluster areas in the business program will be the career performance standards in communication skills, employment literacy, interpersonal skills, occupational safety, personal skills, technological literacy, and thinking and problem-solving skills that are described in box 3-A in chapter 3. The assessment of these career performance standards will be part of the new statewide system of assessment for vocational education that will consist of student portfolios and on-demand assessment (also described in box 3-A).

This core occupational skills model of New York and California is similar to the reorganization of the German apprenticeship system that has occurred over the past 10 years or so to broaden the skill content of the training. A similar branching structure of broad occupational areas that divide into specialties over the 3-year course of an apprenticeship program has been adopted.

DESIGN AND TECHNOLOGY

The fourth approach to defining broad technical skills is to view technology itself as something that is worthwhile for all students to learn about, and not just students who are headed for work or 2-year programs. From this viewpoint, learning about technology as a product of human civilization as well as developing the capability for actually using it to accomplish practical purposes should be part of the education of all students, starting in the early grades. Involving all students in learning about technology and how to use it could help to erase some of the distinctions often made in schools between students headed for college and working with their minds, and those destined for working with their hands. Increasingly, technology itself is blurring this distinction. Teaching technology to all students in this way frames the issue of who needs what technical skills in the broadest possible terms; capabilities

for design and technology are needed to some extent by everyone. At the same time, if started early enough, technology education could help to motivate students toward pursuing technical careers and provide them with foundations of knowledge and capabilities for converting thought into action.[32]

A variety of different forms of technology education exist in the United States and abroad. One version may offer considerable potential for helping all students develop broad technical skills. In Great Britain, this new form of technology education is known as design and technology. It has developed over the past 20 years and is now 1 of the 10 major subjects of the new national curriculum of the British schools.

The significant term is "design." In design and technology, students learn as much or more about becoming a designer as they do about the workings of the technology itself. Becoming a designer involves acquiring three kinds of knowledge and capabilities:

1. procedural knowledge of how to design;
2. capability of communicating complex ideas with clarity, confidence, and skill; and
3. conceptual knowledge and understanding of the use of materials, energy systems, and aesthetics in designing and building technological systems—along with knowledge and awareness of people and their needs.

The procedural knowledge of design involves learning how to discern needs and thoroughly investigate them as an architect or engineer would, prepare a design brief or plan for meeting those needs, generate ideas and develop them into a design, build the design, reflect on the quality of the result in meeting the original needs, and then act on the reflection to produce a better design. The subject is taken by both boys and girls, and both high and low ability students, beginning in the first grade. It is conducted to appeal to all students—those oriented mainly to the humanities, and those more scientifically or technically inclined.

Learning how to design involves learning how to think concretely in complex situations. It involves using knowledge and logical thinking to produce better designs. It also involves making choices and weighing the desirability of alternative designs from social, economic, aesthetic, and human points of view, as well as from the perspective of producability.

A thoroughly developed performance assessment in England has shown the effectiveness of the design and technology curriculum for teaching students how to become designers. Normally, girls perform better than boys on the design tasks of investigation and reflection, while the boys excel at developing designs and building them into final products. Lower ability students tend to be better at design tasks that are tightly structured, and higher ability students tend to be better on more openly structured tasks. However, a clear finding of the assessment is that the more students have been exposed to design and technology, the smaller these differences become.

The assessment also showed the interdependence of thinking and doing in technology. Students given the same set of five assessment tasks but in different order were much better able to identify the strengths and weaknesses of an original design when they had previously been given a task of actually designing an improved product. Providing students with a brief period of discussion midway in the development of a design, where the students described their plans and received peer comments, immensely improved the quality of their final products. The assessments showed no simple relationships between sequences of design steps and thought processes that lead to quality results, and those that do not. In other words, the process of design

[32] This section draws on Richard Kimbell, University of London, Goldsmith's College, "Technology in the School Curriculum," OTA contractor report, October 1993. This contractor report is available from OTA's Education and Human Resources Program.

cannot be reduced to a checklist or a methodical series of steps.

COGNITIVE SKILLS

Finally, a fifth approach to defining broad technical skills is the concept of cognitive skills. This concept comes from cognitive science research on problem solving and trouble shooting in a range of occupations, apprenticeship learning, and learning in academic subjects. Much of the research has been generated by efforts to explain the differences between the cognitive skills of experts and novices.

This research shows that experts draw on tightly integrated structures of procedural and conceptual knowledge. This knowledge is also highly contextualized; it is specific to the particular organization and technological environment of the expert. Over time, this knowledge becomes schematically organized, according to specific problems and situations—it is not abstract. This schematic knowledge enables experts to understand the complex relationships necessary for skilled performance. Experts differ profoundly from novices in the speed with which they can access this knowledge; a novice must try to solve each problem de novo.[33]

Substantial proportions of the knowledge of experts are procedural and tacit; experts know what to do and when, but they cannot always express what they do exactly. Their procedural knowledge, together with the conceptual and contextual knowledge underlies what scientists call metacognition—the ability for goal setting, planning, adaptation, and learning that allow for expert problem solving.[34]

Cognitive skills are acquired in stages. Initially, a skill is heavily dependent on declarative (verbal) knowledge. In the declarative stage, the learner either encounters or is instructed in the facts and procedures relevant to the execution of a particular skill.[35] These facts are stored in memory as statements, which can be verbalized and recalled one-by-one as required. The novice uses this knowledge interpretively—that is, the novice might say: "The situation is 'a,' therefore I should do 'b.' " The novice then would be able to do "b." In this declarative stage, general problem-solving strategies are employed by the novice to organize these steps and bring each one into play until the problem is solved or the goal is reached. An example of such a general, or "weak method," strategy is solving problems by analogy, or mimicking the steps that were followed in successfully solving an apparently similar problem. The strategy is general because it does not depend on knowledge of the area in which it is being applied. "Strong methods" are specific to a domain and are heavily dependent on knowledge of it.

The second stage is "skill automation" or compilation of the cognitive skill. In skill automation, weak methods of problem solving are transformed into strong methods.[36] This occurs through gradual transformation of the factual and conceptual knowledge acquired in the declarative stage into a procedural form that is highly organized so that it can be accessed and used with minimal conscious reasoning activity.[37] The skill becomes "automatic" and the load on working memory of recalling the specific steps involved becomes much less. The skill can be performed

---

[33] Robert Glaser et al., "Implications of Cognitive Psychology for Measuring Job Performance," in Wigdor and Green (eds.) op. cit., footnote 12, pp. 1-2.

[34] Ibid.

[35] James M. Royer et al., "Techniques and Procedures for Assessing Cognitive Skills," *Review of Educational Research*, vol. 63, No. 2, summer 1993, p. 204.

[36] J.R. Anderson, "Skill Acquisition: Compilation of Weak-Methods Problem Solutions," *Psychological Review*, vol. 94, No. 2, 1987, pp. 192-210.

[37] Royer et al., op. cit., footnote 35, p. 206.

without consciously thinking about it and the speed increases. Studies of apprenticeship have shown that this compilation process is aided by: a) situated learning, where students execute tasks and solve problems in a real environment where the purposes of the knowledge are clear; b) external support, or "scaffolding," available from a tutor or master to model the ideal performance, support learning, and supply missing pieces of knowledge; c) fading, or withdrawal of support, as the skill is acquired; and d) learning activities that are carefully sequenced to be both sensitive to student needs and robust enough to foster integration and generalization of the skill.[38]

The third stage is skill refinement, or proceduralization, which is a continuation of the compilation process of the second stage. In this stage, performance of the skills is speeded up by weeding out nonessential steps and strengthening associations between events that may occur in performing the skill and effective actions in response to them.[39] As a result, performance of the skill becomes much more sensitive to small but critical situational differences and the flexibility of response to unexpected situations or new data greatly increases.

This model of skill acquisition has been shown to account for many of the differences in the cognitive skills of novices and experts, and to accurately describe the learning processes through which the skills are acquired.[40] The model has been applied in a wide range of domains from electronics troubleshooting, power plant control, and financial planning to tennis.

The knowledge of experts is thus highly procedural and integrated in nature rather than conceptual and detached from contexts of use, as is so much of the learning process and the knowledge acquired in school. Facts, routines, and concepts are bound together with rules for their application and to conditions under which the knowledge is useful.[41] The knowledge is highly goal oriented and conditional. It includes knowledge of specific goals and interrelated subgoals, methods that are employed in pursuing those subgoals, selection rules for choosing subgoals, evaluating methods based on experience, and prerequisite cognitive skills. For example, the goal of the relatively simple cognitive skill in word processing of moving text from one point in a document to another point can be broken down into a conditional series of subgoals and procedural steps. (To move a block of text from A to B, first block the text to be moved, then choose the "move" operation from the menu, and so forth). Experts have been shown to be much better at identifying the hierarchy of subgoals, or intermediate steps, for accomplishing such tasks than novices. In the language of cognitive science, experts are much better than novices at mapping the "problem space" of procedures to employ in solving a problem or generating an idea and deciding on the order. Novices, for example, may not be able to identify the problem space or structure of goals, and simply try the few procedures they know one by one to see which one may work. Selection rules refer to the knowledge acquired by the expert of the particular conditions under which specific procedures should be employed. Prerequisite skills could be, for example, capabilities for performing all of the measurements needed in executing a procedure.

Cognitive research has shown that the organization of the knowledge structures of experts is what explains the flexibility and speed with which they are able to access knowledge and use it. One of the profound aspects of the knowledge organization of experts is their "depth of problem

---

[38] Sherrie P. Gott, "Apprenticeship Instruction for Real-World Tasks: The Coordination of Procedures, Mental Models, and Strategies," *Review of Research in Education*, vol. 15, 1988, p. 99.

[39] Royer et al., op. cit., footnote 35, p. 206.

[40] Glaser et al., op. cit., footnote 33.

[41] Ibid., p. 7.

representation." Using their stored knowledge, experts are able to rapidly induce principles from the given features of problems and then represent the problems in terms of these principles. This gives them quick access to patterns of events and associated sequences of preferred moves and procedures. They are able to do this because they do not have to first load this knowledge into their working memory before they are able to access it and use it. The problem representations of novices, on the other hand, tend to be expressed in terms of literal objects, the surface features of problems, or the events given explicitly in a problem statement.[42]

One example of the depth of problem representation of experts is "chunking." For example, expert electronics technicians have been shown to be much better at reproducing circuit diagrams flashed before them for a brief period of time than novices, but only when the circuits are correct. When the circuits are wrong, there is no difference between experts and novices in their ability to reproduce the diagrams.[43] Skilled electronics technicians also tend to reconstruct symbolic drawings of circuit diagrams they are working with according to the functional nature of the elements in the circuit, while novices tend to reproduce the circuits based on the spatial proximity of the elements. The "chunks" of knowledge they draw on to produce these diagrams tend not to reflect knowledge of function.[44]

In addition to the contextual and procedural knowledge in the knowledge structures of experts is a third component, the capability for visualizing or mentally modeling the features or operation of the technological devices or systems, or representing the problem space of interpretations and conceivable actions. These mental models provide essential connections between the procedural and contextual knowledge, enabling the expert to relate knowledge of one form to the other and build representations of the situations at hand.[45] These visual representations help to overcome the limiting serial nature of language and procedural knowledge to provide more adequate means of explaining or interpreting complex and/or dynamic phenomena encountered.[46] An important finding in studies of the diagnosis of x-rays by physicians, for example, shows that experts have a reservoir of preexisting schemata in their heads of normal and abnormal configurations of the organs or tissues that are triggered early in a diagnosis.[47] The diagnosis involves fitting the schemata to the present case until the most consistent explanation is found.

Mental models are essential for effective causal reasoning about the sources of problems and potential solutions. Studies of experts have shown how much better they are than novices at visualizing the systems with which they are working at different levels of abstraction and working back and forth across these levels of abstraction to invoke the level of analysis most appropriate for the diagnosis or task at hand.[48] For example, in diagnosing a broken piece of equipment, experts may employ their schematic knowledge of the functions of the equipment and its components, factual knowledge about the operational characteristics of the individual components, and orderings of how observed symptoms of one kind are conditionally related to others. Much of this reasoning about problems tends to be qualitative

---

[42] Royer et al., op. cit., footnote 35, pp. 217-222.

[43] Ibid., p. 218.

[44] Robert Glaser, "Expertise and Assessment," *Cognition and Testing*, M.C. Wittrock and E.L. Baker (eds.) (Englewood Cliffs, NJ: Prentice Hall, 1991), pp. 17-30.

[45] Glaser et al., op. cit., footnote 33, p. 3.

[46] Gott, op. cit., footnote 38, p. 123.

[47] Glaser et al., op. cit., footnote 33, p. 9.

[48] Gott, op. cit., footnote 38, p. 124.

in nature rather than quantitative (e.g., "if the signal is present there, then it should be present here"). Overemphasis on quantitative reasoning in the early stages of acquiring skills has consistently been shown to inhibit understanding of the causal principles involved.[49] For example, students in the physical sciences who have mainly studied in the traditional modes of highly quantitatively oriented instruction have consistently shown deficiencies and mistakes in their understanding of the underlying causal principles in a domain.[50]

A number of experimental efforts have been undertaken to build computer-based tutors for training experts based on these concepts of the nature and acquisition of cognitive skills. Evaluation of the effectiveness of these tutors has shown the strongest results for systems that coordinate the teaching of contextual and procedural knowledge so that the uses of visualization and modeling in problem solving are made explicit.[51] Tutors that mainly support the acquisition of procedural knowledge yield trainees who are less able to deal with complexities and uncertainty. Such training has usually been found to be even less effective and efficient than methods of informal observation on the job. A related finding is that the presentation of "... abstract representations of device interdependencies for learners to inspect and manipulate seems central to the development of envisioning capabilities."[52]

This model of the acquisition of cognitive skills has important implications for broad technical skills. The cognitive model supports the job task and procedural orientation of job competency approaches over attempting to teach or assess general abilities directly, as implied by the vocational aptitude and core occupational skills approaches. At the same time, the model implies that the limitation of learning to the rote memorization of procedures within a fixed and narrow task environment without the simultaneous introduction of concepts and support for modeling and visualization will inhibit the development of expertise and result in trainees who are not capable of responding flexibly to new problems. The model therefore suggests that the teaching and assessment of broad technical skills should include methods of eliciting respondents' capabilities for visualizing and conceptualizing problems in context, and not just recalling facts from memory even if the facts are closely related to those contexts and demonstrate mastery of procedures. The model does not support decontextualized approaches to the assessment of broad technical skills, like those employed in traditional testing for vocational aptitudes or multiple abilities, and especially not at expert levels of performance. The decision orientation of the model strongly suggests that the assessment of broad technical skills should include strategies for assessing metacognition—the capability to choose among goals and methods and alternative interpretations of evidence in different contexts.

The cognitive model also strongly suggests that assessment should focus on both the procedural and conceptual aspects of broad technical skills in authentic contexts, rather than the recall of isolated bits of knowledge on demand. Reliance should be placed on using multiple methods of assessment rather than written testing, although both may be needed. The highly integrated and contextualized nature of the process of knowledge compilation lying at the heart of the cognitive model implies that learning broad technical skills must also be active, so that students are constructing knowledge for them-

[49] Ibid., p. 127.

[50] Ibid. Also see Howard Gardner, *The Unschooled Mind: How Children Think, and How Schools Should Teach* (New York, NY: Basic Books, 1991).

[51] Gott, op. cit., footnote 38, p. 161.

[52] Ibid.

selves, gaining facility in integrating it from different sources, and using it for practical ends.

Finally, the cognitive model raises a caution flag against the possibilities of broadening skill definitions and methods of assessing and certifying them simply through expanding indefinitely the job and task environments in which they apply. The reason for this skepticism is the prominent role of contextual knowledge in the development of expertise, and the highly organized and tightly coupled ways in which the knowledge of experts is structured.

A different model of how to achieve breadth of skill is suggested by the cognitive model in which breadth is achieved through depth. The view is that the best way to achieve breadth of skill in the long run is through deepening knowledge development within a limited number of areas of specialization, rather than attempting to take on ever wider circles of problems. The concept is achieving "breadth through depth" rather than breadth alone. The problem with the breadth through depth approach is that expertise developed through specialization and depth of pursuit can be "brittle;" it does not transfer well to new situations and new kinds of problem environments. When confronted with new problems in unfamiliar areas, the narrowly trained learner can be immobilized and incapable of responding productively. In order to achieve breadth, the strategy of developing and assessing depth must be coupled with considerable experience in dealing with new problems and novel situations.

Cognitive theory is not very helpful so far in prescribing exactly how transfer occurs and beyond that how learning should occur in order to maximize capacity for transfer, except to say that the knowledge and skills acquired will be brittle if learners are not confronted with novel problems to solve within their domains of expertise and in related domains. In this light, breadth of skill becomes the individual's capability for using knowledge from domains in which one is expert in new domains. Whether knowledge and skills from domains in which one is expert are applied in the new domains as "finite elements" or instead facilitate the acquisition of knowledge and skills in the new areas is an unresolved issue.

FINAL NOTE

The aptitude model suggests that the ability to transfer expertise in novel situations and to new domains needs to be defined in terms of both demonstrated capacity for responding cognitively to new situations and evidence of broader general aptitudes that strongly relate to the ability to transfer and learn in new situations. This melding of two approaches is also evident in the addition of categories of knowledge and skill to the broad occupational tasks that are identified in the approach of occupational mapping.

In the occupational mapping approach, breadth of technical skill is described through identifying tasks presumed to define competencies that cut across a range of occupations and form the core of expertise. Their manifestations may be somewhat different in different job contexts and industry segments but in outline they are similar. The critical role of contextual knowledge in expertise, as described in the section on the cognitive model, indicates that it remains an open question whether individuals who are found to be competent in one job context according to the occupational mapping approach would be able to demonstrate equal levels of the same competencies in other job contexts.

These five approaches to defining broad technical skills provide starting points for moving to concepts that are at once deeper, more unifying, and more specific than the concepts underlying the various approaches. Another need is to become clearer about the content of technical skills and how they differ from other capabilities, such as interpersonal and communications skills, motivation, and academically oriented competence.

Implementation of Performance Standards and Measures

6

Under the 1990 Perkins Act Amendments, Public Law 101-392, the primary responsibility for planning and implementing vocational education standards and measures rested with the states. Federal law established only minimum content and process requirements for statewide systems, leaving key decisions to the discretion of state boards.[1] This decentralization of authority allowed for variation in state systems, consistent with the reality that states differed greatly in their structures for vocational education and in their expertise in performance-based accountability.[2] Further, it reflected the absence of national consensus about standards for evaluating vocational education.

In keeping with this decentralization, federal administrative responsibilities regarding performance standards are limited but still quite important. The amended Perkins Act charged the U.S. Department of Education (ED) with providing technical assistance to states for the development of accountability systems (section 115) and with conducting or overseeing research on standards and measures (sections 402, 403, 404, and 422). Furnishing guidance on performance standards was also an implicit part of ED's obligation to issue regulations for the Perkins

[1] Among the decisions left to states were which standards and measures would be used, how local program quality would be judged, and what type of assistance would be provided to programs making insufficient progress toward standards.

[2] A national survey conducted in 1991, prior to the Perkins Act deadline for developing performance systems, found that about one-half of the states had used specific performance standards or measures for vocational education in the past. E. Gareth Hoachlander and Mikala L. Rahn, *Performance Measures and Standards for Vocational Education: 1991 Survey Results* (Berkeley, CA: National Center for Research in Vocational Education, 1992), p. 4.

Act and to conduct regular oversight and monitoring of federal vocational education programs. Administering a federal competitive grant program to develop national skill standards for particular industries and trades (section 416) is another related component of ED's role.

Some of these responsibilities are being carried out directly by ED's Office of Vocational and Adult Education (OVAE); some are being conducted through other entities, particularly the federally supported National Center for Research in Vocational Education (NCRVE).

TECHNICAL ASSISTANCE EFFORTS

Providing technical assistance is one of ED's most critical duties concerning performance standards. The move toward performance-based accountability is a pivotal change for vocational education. In developing their systems, states—especially those with little prior experience in performance-based evaluation—were likely to encounter several difficult and highly technical issues for which additional guidance would be helpful. Recognizing the complexity of the task, the law gave states 2 years, until September 25, 1992, to put in place their new systems and charged ED with furnishing technical assistance.

ED has delegated much of the responsibility for providing technical assistance to NCRVE. For several reasons, NCRVE was a logical choice for this assignment. Research, dissemination, and training on issues of performance standards were already part of the center's mission under the revised Perkins Act. In addition, ED concluded that the center was more likely than OVAE to offer the necessary expertise in technical issues of evaluation, testing, and measurement, especially given OVAE's current staffing levels.[3]

Under the present arrangement, OVAE handles day-to-day communications with states and oversees implementation of state accountability plans. For additional guidance on performance and evaluation issues, OVAE often refers states to NCRVE.[4] In providing technical assistance, NCRVE has undertaken five special efforts to help states implement the new accountability requirements:

- three regional workshops held in February and March of 1992, and one national technical assistance session held in July 1992, all jointly sponsored with OVAE;[5]
- operation of a clearinghouse;
- publication of a practitioner's guide on accountability;[6]
- research studies on the implementation of state standards and measures; and
- a national conference on performance standards in vocational education in the summer of 1993.

NCRVE has received some special funding to help carry out these activities. From an additional $2 million appropriated by Congress to NCRVE, ED earmarked approximately $200,000 for technical assistance on performance standards.

The core of NCRVE's initial technical assistance efforts was the 1992 workshop series; about two-thirds of the $200,000 supported this activity. During the workshops, state officials responsible for developing the new accountability systems had the opportunity to share practices and receive expert advice on issues related to stan-

[3] Joseph Casello, branch chief, Program Analysis Branch, Office of Vocational and Adult Education, U.S. Department of Education, personal communication, June 2, 1993; E. Gareth Hoachlander, director, Planning and Evaluation, National Center for Research in Vocational Education, personal communication, June 1, 1993; and Debra Nolan, project director, Business and Education Standards Program, Division of National Programs, Office of Vocational and Adult Education, U.S. Department of Education, personal communication, June 9, 1993.

[4] Casello, op. cit., footnote 3.

[5] Regional workshops of 2+ days each were held in Washington, DC, St. Louis, and San Francisco; the national workshop was held in Minneapolis.

[6] E. Gareth Hoachlander et al., *Accountability for Vocational Education: A Practitioner's Guide* (Berkeley, CA: National Center for Research in Vocational Education, 1992).

dards, measures, and assessments. OVAE staff participated in all the workshops and were members of the NCRVE advisory group charged with planning the workshops and reviewing NCRVE materials regarding performance standards.

The remainder of the $200,000 has helped support a clearinghouse, which gathers current information about state accountability plans and provides telephone technical assistance.

Through its *Practitioner's Guide* and other materials, NCRVE has amplified the limited guidance contained in the law and regulations with detailed suggestions, examples, and recommendations. Among them are concrete examples of standards, learning measures, labor market outcomes, assessment methods, and strategies for evaluating access of special populations.

NCRVE also identified six basic features that should be incorporated into a well-designed accountability system and offered expert opinions on several key issues. For example, NCRVE recommends that state accountability systems go beyond the 2 minimum measures required by law and include from 6 to 12 measures.[7] The center also suggests that student performance be assessed both in terms of gains over time and absolute attainment, so that growth by students who began at a low level is not overlooked.[8] NCRVE has further encouraged states to phase in elements of a more ambitious system over time and to continue monitoring and modifying their accountability systems as they gain experience.[9]

To meet the need for more sophisticated guidance on technical issues, both NCRVE and OVAE plan to continue offering technical assistance, maintaining the clearinghouse, and conducting workshops and training session. A follow-up conference jointly sponsored by NCRVE and OVAE was conducted in July 1993. The purpose of the meeting was to provide state administrators with a forum to share experiences and help ED determine what additional technical assistance is needed.[10]

Participants in the meeting identified the main issues where further technical assistance is needed by the states in implementing performance standards. Discussion focused on: 1) using the information that will be generated in local program reviews for purposes of improvement, and 2) what program improvement plans should be like in order to be really helpful. Many states are also looking for assistance in how to set standards using rigorous methods. Concerns were expressed that not enough is known about the reliability and validity of the skill standards and methods of assessment that are being used. Issues of reliability and validity will grow in importance as performance standards are implemented.

As yet, no technical assistance efforts have been directed specifically toward testing and assessment for the implementation of performance standards.

ED REGULATIONS AND MONITORING

Issuing regulations and monitoring state compliance for federal vocational programs are among the major administrative responsibilities of ED, and both affect state implementation of performance standards.

Evaluation and accountability requirements became a highly controversial issue during the development of regulations for the Perkins Act.[11] In October and November of 1990, ED held regional meetings, as required by section 504 of the Perkins Act, to obtain public input prior to

[7] Ibid., p. 7.

[8] Ibid., p. 13.

[9] Ibid., p. 112.

[10] Casello, op. cit., footnote 3.

[11] It is noteworthy that evaluation, standards, and measures were not particularly controversial during congressional consideration, overshadowed by other heated issues, such as allocation formulas and the distribution of funds between secondary and postsecondary programs.

drafting regulations.[12] Following these meetings, the Secretary, as further required by law, selected three major issues to undergo the process of negotiated rulemaking, which provides state and local representatives with a forum to discuss and make recommendations to the Department on issues of complexity, contentiousness, or substantial consequence. One of the issues chosen was whether state and local vocational agencies should apply program evaluations to all of their vocational education programs or only to the federally funded components or projects.[13]

On October 11, 1991, ED published proposed regulations that took the position that local entities receiving Perkins basic grant funding had to evaluate the effectiveness of *all* their vocational education programs, not just those receiving Federal funding.[14] A surge of public comment opposed this interpretation, many arguing that it was unduly costly and burdensome, that it was inconsistent with the language in the law, or that it represented unwarranted federal intrusion.[15]

Final regulations were not issued until August 14, 1992—less than 6 weeks before the state standards and measures were due, and well past the deadline for ED regulations specified in the General Education Provisions Act.[16] Debate within the administration over the evaluation issue seems to have been the main reason behind the delay.[17] Thus, states were compelled to proceed with development of standards and measures in the absence of a definitive interpretation about their scope.

The final regulations narrowed the evaluation to encompass only "... the particular projects, services, and activities ..." receiving Perkins Act basic or special funding, unless the state determined that a broader evaluation was needed.[18] The final rules also contained other important clarifications. When conducting evaluations, local vocational agencies could rely on representative sampling techniques. In addition, grantees could use Perkins funds to pay for mandated evaluations, without regard for the 5 percent ceiling on local administrative costs.[19]

Mirroring the legislation, the regulations avoided further specificity on such issues as what "competency gains" mean and how they should be measured, whether separate standards for secondary and postsecondary programs are encouraged or desirable, how special populations should be addressed, and how basic and advanced academic skills should be measured. These and other specialized issues were left for states to decide, with advice from NCRVE and other sources.

According to OVAE officials, every state appears to have met the minimal requirements for a statewide performance system; many states plan to phase in or expand the components of the system over time. As a next step, OVAE is currently conducting onsite reviews of state implementation.[20]

---

[12] Meetings were held in Philadelphia, Atlanta, Kansas City, and San Francisco.

[13] Paul M. Irwin and Richard N. Apling, *Vocational Education: Major Provisions of the 1990 Amendments (P.L. 101-392)* (Washington, DC: Congressional Research Service, 1991), p. 18.

[14] 56 *Federal Register* 51448 (Oct. 11, 1991).

[15] 57 *Federal Register* 36841 (Aug. 14, 1992).

[16] Section 431(g) of the General Education Provisions Act (20 U.S.C. 1232) requires final regulations for Department of Education programs to be promulgated within 240 days of enactment of the applicable authorizing statute, unless a revised schedule is approved by the congressional authorizing committees. It is not uncommon, however, for the Department to have missed this deadline.

[17] John F. Jennings, counsel, House Committee on Education and Labor, personal communication, Feb. 23, 1993; and "ED Set to Issue Voc Ed Rules After Ford Increases Heat," *Education Daily*, Aug. 12, 1992, p. 3.

[18] 45 CFR 403.191 and 403.201.

[19] 45 CFR 403.191.

[20] Casello, op. cit., footnote 3.

RESEARCH AND DEMONSTRATION

Several federally sponsored studies are under way that address the issues of performance standards, skill standards, and accountability.

As required by law, NCRVE is conducting research on performance standards and accountability systems. Findings of a baseline survey of prior state experience with vocational education measures and standards were published in March 1992.[21] A second study describes and analyzes the state accountability systems to date.[22] A third study, conducted jointly with the Rand Corp. and scheduled for completion in 1994, is examining the effects of performance accountability systems on state and local policy.

OVAE has also awarded a contract to study the performance standards adopted by the states, as called for in Section 115(f) of the 1990 amendments. The purpose of the study is to evaluate and review the systems developed by the states. Through a mail survey and case studies in nine states, the study will:

- describe, in some detail, the status of each state's system of performance standards and measures,
- assess the reasonableness and appropriateness of performance standards and measures for specific populations, and
- examine the comparability of the performance standards across states to determine the feasibility of establishing a common core of indicators.

The study is being performed by the Battelle Human Affairs Research Centers.[23]

BUSINESS AND EDUCATION STANDARDS

Another federal activity with relevance to the new Perkins accountability requirements is the joint Department of Education-Department of Labor (DOL) initiative to develop voluntary, national skill standards in various industries and trades. The ED component, called the Business and Education Standards Program, is authorized by section 416 of the Perkins Act. The DOL component, called the Voluntary Skill Standards and Certification Program, is being supported with departmental discretionary funds.[24] Both Departments have made competitive grants to projects to organize and operate business-labor-education partnerships, which will in turn develop national skill standards and certifications for competencies in industries and trades.

Eligible grantees under the ED program include industrial trade associations, labor organizations, national joint apprenticeship committees, and comparable national organizations. The standards developed through ED grants must address at least:

1. major divisions or specialty areas within occupations;
2. minimum hours of study to become competent;
3. minimum tools and equipment required;
4. minimum qualifications for instructional staff; and
5. minimum tasks to be included in a course of study.[25]

The ED program was funded with appropriations of $3.5 million annually from fiscal years 1991 and 1992. DOL reserved $1.2 million from

---

[21] Hoachlander and Rahn, op. cit., footnote 2.

[22] Mikala Rahn et al., *State Systems for Accountability in Vocational Education*, MDS-491 (Berkeley, CA: National Center for Research in Vocational Education, December 1992).

[23] Battelle Human Affairs Research Centers, "Performance Standards and Measures. Evaluation Study," report being prepared for the U.S. Department of Education, Office of Vocational and Adult Education, in progress.

[24] Michaela Meehan, policy analyst, U.S. Department of Labor, personal communication, June 15, 1993.

[25] 57 *Federal Register* 45146 (Sept. 30, 1992).

fiscal year 1992 discretionary funds. Grantees must match federal funds dollar for dollar.[26]

In the first round of grants, awarded in September 1992, ED supported seven projects and DOL six. In the second competition, ED awarded nine new projects.[27] The grant period for the ED projects is 18 months, with one continuation, and the grant period for DOL projects is 1 year.[28]

Both ED and DOL recommend that standards be benchmarked to world-class levels of industry performance and "... tied to measurable, performance-based outcomes that can be readily assessed."[29] The ultimate aim is for the projects to yield standards that could be adopted and used by education and training institutions, labor organizations, business and industry, employers, individuals, and government.

From DOL's perspective, it is particularly critical that the standards developed have support from business and industry; if this occurs, then education institutions and other communities are likely to follow suit.[30]

How the skill standards might mesh with state accountability systems required by the Perkins Act or with Job Training Partnership Act (JTPA) performance standards is a critical issue. According to the ED Business and Education program director, she and the OVAE administrators of the state vocational programs consult regularly, and ED is taking several steps to ensure that the two efforts are coordinated. Regulations and guidance for the Business and Education program emphasize the importance of dissemination and adoption of standards by state officials and other entities. In addition, ED has encouraged grantees to examine existing standards, including those developed by the state technical committees authorized in the original Perkins Act, when developing national standards. Further, almost every partnership includes involvement of state vocational education directors.[31]

The DOL program administrator, however, emphasized that the ultimate goal of the DOL standards program is to produce *national* standards and overcome the fragmentation that could occur if each state proceeds with its own standards. Over the long term, DOL also hopes to be able to use national skill standards to evaluate outcomes for JTPA and other Department programs.[32]

GOALS 2000: EDUCATE AMERICA ACT

Progress toward national skill standards may be further stimulated if the Goals 2000: Educate America Act—a primary education initiative of the Clinton Administration—is enacted.[33] ED and DOL have consistently stressed the relevance of their respective skill standards programs to National Education Goal No. 5 (adult literacy, citizenship, and ability to compete in the workplace). This link would be solidified by Title IV of Goals 2000, which is currently being

---

[26] Debra Nolan, U.S. Department of Education, "Project Abstracts for the Business and Education Standards Program," unpublished report, 1992.

[27] First-round projects will develop standards in the following fields: health sciences and technology; electronics; computer-aided drafting; air conditioning and refrigeration; biotechnical sciences; printing; automotive technicians; industrial laundering; tourism, travel, and hospitality; metalworking; electrical construction; and retail trade. Second-round projects focus on human services occupations, heavy highway construction and demolition, chemical process industries, hazardous materials management, photonics, agriscience, welding, forest and wood production and manufacturing, and food marketing.

[28] A decision is pending on whether additional DOL discretionary money will be provided to continue the existing projects for another year and to make a new round of grants.

[29] 57 *Federal Register* 9490 (Mar. 18, 1992).

[30] Meehan, op. cit., footnote 24.

[31] Nolan, op. cit., footnote 3.

[32] Meehan, op. cit., footnote 24.

[33] H.R. 1804 and S. 846.

considered by Congress. The bill would establish a national skill standards board, charged with identifying broad occupational clusters and encouraging the development of voluntary, national skill standards for each cluster. The legislation authorizes $15 million annually for this purpose.

CONCLUSION

The Department of Education took seriously its mandate to help states implement new accountability systems for vocational education and, with the help of NCRVE and other entities, has provided a reasonable level of technical support on this issue, especially considering the limited staffing capacity of OVAE and the complexity of the issues involved. It is likely that in the future more attention will be required to establish validity and reliability of the methods of testing and assessment being employed in vocational education.

The relationship among state performance standards, the ED and DOL skill standards programs, and new legislation affecting national skill standards raises important issues warranting continued attention.

Appendix A: Legislative Milestones Influencing Accountability in Federal Vocational Education Programs

A

1917: Smith-Hughes Act

- Required states to adopt minimum standards for teacher qualifications, plant and equipment, and hours of instruction in federally supported vocational programs.
- Required states to develop plans with approval of a Federal Board for Vocational Education.
- Identified the outcome of vocational education programs as preparing students to be fit for useful employment.

1963: Vocational Education Act

- Required states to conduct and use periodic evaluations of state and local vocational education programs in light of current and projected "manpower" needs and job opportunities.
- Reserved 3 percent of state grants for ancillary services, including program evaluation.
- Established an ad hoc federal advisory council to review and recommend improvements in vocational education programs.
- Revised state plan provisions to require more detailed assurances and descriptions.

1968: Vocational Education Amendments

- Continued requirements for evaluations based on labor market needs.
- Authorized funds for evaluation and other activities related to state administration.
- Limited federal funding only to vocational programs that could be "... demonstrated to prepare students for employment or for successful completion of such a program or be of significant assistance to participants in making an informed and meaningful occupational choice."
- Created a standing national advisory council to evaluate program effectiveness.
- Established state advisory councils with program evaluation duties.
- Required states to develop both long-range and annual plans.
- Mandated local plans.

1976: Vocational Education Amendments

- Directed states to evaluate the effectiveness of each local program within the 5-year duration of the state plan.
- Required programs that sought to impart entry-level job skills to be evaluated according to the extent to which program completers and leavers found employment in occupations related to their training (or pursued additional education) and were considered by employers to be well trained.
- Expanded the evaluation duties and enhanced the independence of state advisory councils.
- Created local advisory councils.
- Authorized technical assistance and studies from a national advisory council and a new national research center, and authorized a national data and occupational information system.

- Clarified and strengthened the evaluation and monitoring responsibilities of the Department of Health, Education, and Welfare.
- Expanded state planning duties.
- Required states to summarize evaluation findings in an annual accountability report and to assess current and future job needs within the state.

1982: Job Training Partnership Act

- Pioneered use of performance standards to select service providers, encourage quality and efficiency, and target programs for rewards or sanctions.

1984: Carl D. Perkins Vocational Education Act

- Charged states with developing measures of program effectiveness in such areas as occupational skills and basic employment competencies, with separate measures for disabled individuals.
- Required states to evaluate not less than 20 percent of local recipients in each fiscal year.
- Continued requirements for advisory council evaluations, national data on completers and leavers, state needs assessments, and a national assessment of vocational education.
- Deleted requirement for states to evaluate job placement and employer satisfaction for program completers and leavers.
- Simplified state planning and reporting requirements.
- Established state technical committees to develop skill inventories for specific occupations.

1988: Hawkins-Stafford Elementary and Secondary School Improvement Amendments

- Amended the Chapter 1 program for disadvantaged students to require schools that did not meet standards for increased achievement among participating children to develop and implement "program improvement" plans.

1988: Family Support Act

- Created the Job Opportunities and Basic Skills (JOBS) program for welfare recipients, which required the Secretary of Health and Human Services to develop outcome standards for judging individual progress by October 1, 1993.

1988: Hunger Prevention Act

- Amended the Food Stamp Employment and Training Program to require the Secretary of Agriculture to establish performance standards for measuring the employment outcomes of participants.

1990: Carl D. Perkins Vocational and Applied Technology Education Amendments

- Required states to develop standards and measures of performance for vocational education programs, reflecting attainment of job or academic skills, school retention and completion, or placement.
- Required local recipients to: conduct an annual evaluation of vocational programs based on state standards and measures; review the progress of special populations; and assess student knowledge about the industry they are preparing to enter.
- Required local recipients that did not show substantial progress to develop and implement program improvement plans.
- Directed states to develop measurable goals and accountability measures for special populations and required local recipients to describe how access to programs of high quality will be provided for special populations.
- Required states to conduct an initial assessment of program quality and needs using measurable objective criteria.
- Authorized technical assistance, advice, and data collection on performance standards and evaluation by the Department of Education, the National Center for Research on Vocational Education, the National Occupational Information Coordinating Committee, and state advisory councils.
- Authorized grants to business-education committees to develop skill standards in trades and industries.

Appendix B: List of Acronyms

B

ACT — American College Testing
AEA — American Electronics Association
ASVAB — Armed Services Vocational Aptitude Battery
ATIB — Academic Test Item Bank (Arizona)
C-TAP — Career-Technical Assessment Project (California)
DOL — U.S. Department of Labor
ED — U.S. Department of Education
FBVE — Federal Board for Vocational Education
IOCT — Industrial Occupational Competency Testing
JOBS — Job Opportunities and Basic Skills
JTPA — Job Training Partnership Act
NCRVE — National Center for Research in Vocational Education
NIE — National Institute of Education
NOCTI — National Occupational Competency Testing Institute
OOTC — Oklahoma Occupational Testing Center
OVAE — Office of Vocational and Adult Education (U.S. Department of Education)
OVCAP — Ohio Vocational Competency Assessment Program
R&D — research and development
SCANS — Secretary's Commission on Achieving Necessary Skills
SOCAT — Student Occupational Competency Achievement Testing
TABE — Test of Adult Basic Skills
TIERS — Title I Evaluation of Reporting System
TOCT — Teacher Occupational Competency Testing
V-TECS — Vocational-Technical Education Consortium of the States

Index

Academic skills
  changes in skills assessed, 16, 61-62
  definition, 2, 3, 46
  norm-referenced assessment, 7, 9, 16, 39, 42
  sources of skill information, 13-14, 48
  state assessment policies, 12, 46, 63-64
  statewide academic testing programs, 39, 42, 63
  strategies for obtaining assessment information, 13-14, 64
Academic Test Item Bank, 49-50
Accountability
  changing themes, 4-6, 32-34
  current regulations, 22-26
  legislative history, 4-6, 26-32
  mechanisms, 32-33
  performance-based, 4-5, 21-34
  requirements, 2, 33-34
  special populations, 3, 24
Acronym list, 114
ACT. *See* American College Testing
AEA. *See* American Electronics Association
AFQT. *See* Armed Forces Qualification Test
American College Testing, 10, 67
  assessment development process, 70
  job profiling, 71-72
American Electronics Association, 18, 91-94
American Society for Training and Development, 75
American Vocational Association, 75
Annual review requirement, 5
Applied mathematics test, 69
Applied technology test, 69
Aptitude tests, 84-85
Arizona, 49, 53
Armed Forces Qualification Test, 90
Armed Services Vocational Aptitude Battery, 17, 86-91
Assessment. *See also specific skills assessed*
  definition, 7, 36-37, 39
  differences in state policies, 11-12, 41-42
  methods, 7, 39
  origins of assessment practices, 7, 38-41
  resources, 2, 9-11, 42-44
  state policies, 2, 11-12, 24-25, 41-42
  use of results, 13, 51-52
ASVAB. *See* Armed Services Vocational Aptitude Battery
ATIB. *See* Academic Test Item Bank

Broad technical skills, 44, 65
  changes in skills assessed, 15-16, 65
  cognitive skills, 19-20, 97-101
  core occupational skills, 18, 94-95
  definition, 2, 3, 16, 44, 46, 84
  design and technology, 18-19, 95-97
  lengthening the skill horizon, 36, 83
  occupational maps, 18, 91-94
  state assessment policies, 13, 65
  vocational aptitudes, 17, 84-91
Business and education standards, 107-108. *See also* Industry skill standards

C-TAP. *See* California Career-Technical Assessment Project
California, 53, 59, 95
California Career-Technical Assessment Project, 59, 60-61
Carl D. Perkins Act. *See* Perkins Act of 1984
Certification of student competence for employers, 2, 5, 13, 51
CETA. *See* Comprehensive Employment Training Act
Cognitive skills, 19-20, 97-101
Competency assessment
  definition, 8
  methods, 39-41
  shift away from, 15, 57-58
Competency testing
  definition, 8, 39
  item banks, 43, 74-75, 75-76
  shift toward, 14-15, 57-58
  written, 39

Competitive events, 8, 39
Comprehensive Employment Training Act, 34
Computer-based tutors for training experts, 100
Core occupational skills, 18, 94-95
Criterion-referenced assessment, 7, 16, 43, 49, 63
  definition, 9, 39
  NOCTI, 77
  VTECS, 73
  Work Keys, 69-70
Curriculum
  materials, 11-12, 43-44, 72, 73
  where skills are assessed, 48-51, 54

DACUM process, 77
Delaware, 78
Design and technology, 18-19, 95-97
Disabled students, 24
DOL. *See* U.S. Department of Labor

ED. *See* U.S. Department of Education
Effectiveness of vocational programs
  local reviews of, 1, 13, 21, 24, 27, 32
  levels of skills to be achieved, 4
  need for improvement, 4-5, 6
  occupations to be trained for, 4
  outcome measures, 5
Elementary and Secondary Education Act of 1965, 31
Exit exams, 13-14, 48-50, 63
Expansion of testing and assessment programs, 14-17, 55-65

Family Support Act of 1988, 112
FBVE. *See* Federal Board for Vocational Education
Federal Board for Vocational Education, 27
Food Stamp Education and Training program, 34

General Education Provisions Act, 106
Generic workplace skills, 45
  changes in skills assessed, 16, 64-65
  definition, 2, 3, 46, 84
  state assessment policies, 13, 46
Glossary of testing terms, 8
Goals 2000: Educate America Act, 45-46, 108-109
Grade levels of academic and vocational testing, 16, 42, 48, 50
Great Britain, 96
GT. *See* General Technical

Hawkins-Stafford Elementary and Secondary School Improvement Act
  Amendments, 112
Hunger Prevention Act of 1988, 112

Industrial Occupational Competency Testing, 76
Industry skill standards projects, 17, 23
  demonstration projects, 18, 25, 91
  purposes of federal demonstrations, 107
  state directors' study, 3, 38
  skill standards board, 45, 109
  technical committees and, 4, 30
Instructional diagnosis, 2, 8, 11, 13, 30

IOCT. *See* Industrial Occupational Competency Testing
Iowa, 56
Item banks, 43
  Arizona Test Item Bank, 49
  state resource, 11-12, 37, 43, 44
  use and availability, 38, 55, 82
  VTECS, 18, 74-76, 81

Job Opportunities and Basic Skills training program for welfare recipients, 22-23, 34
Job placement
  and performance standards, 5, 24, 25, 42
  evidence of, 23, 28-29
  JTPA, Title II, 30-31
  evolution of legislation, 32
Job profiling, 71-72
Job Training Partnership Act of 1982, 22-23, 25, 30-31, 34
JOBS. *See* Job Opportunities and Basic Skills training program for welfare recipients
JTPA. *See* Job Training Partnership Act

Kansas, 56

Learning or competency gains
  and performance standards, 5, 24, 35, 105, 106
  difficulty of measuring, 16-17, 50, 63
  Chapter I, 32
  NOCTI, 79
  Oklahoma testing program, 42
  Work Keys, 72
Learning test, 70
Legislation. *See specific laws*
Listening and writing test, 69
Location of information test, 69

Marine Corps Job Performance Measurement Project, 89
Math. *See* Applied mathematics test
Methodology
  state survey, 2-4, 37-38
Military Entrance Processing Command, 86
Mississippi State University, 3, 38
Motivation tests, 70, 102
MPR Associates, 3

National Academy of Sciences, 87, 89
National Assessment of Vocational Education, 22
National Association of State Directors of Vocational Education, 3

National Center for Research in Vocational Education, 35-36, 104
National Governors' Association, 23
National Occupational Competency Testing Institute, 9-11, 43, 67, 76
National Skills Standards Board, 45-46. *See also* Industry skills standards
National system of performance standards. *See* Performance standards
NCRVE. *See* National Center for Research in Vocational Education
New Jersey, 5
New York, 41, 94, 95
NOCTI. *See* National Occupational Competency Testing Institute

Observation test, 70
Occupational assessment, 3, 16
  form of measurement, 53-55
  shift to written testing, 14, 58
  skills assessed, 44-47, 59-61
  state assessment policies, 12, 41
  state resources, 42-44
  why skills are assessed, 51-53
Occupational maps, 18, 91-94
Occupational skills. *See* Core occupational skills
  changes in skills assessed, 55-59
  definition, 2, 46
  differences compared to academic skills, 13
  shift to written testing, 15
  sources of skill information, 13-14, 16, 48-49
  state assessment policies, 12-13, 35, 41
  state testing resources for, 43
  why skills are assessed, 13-14, 51-53
Office of Technology Assessment, 1-2, 7, 11, 17, 36, 83
Office of Vocational and Adult Education, 104, 107
Ohio, 14, 56, 82
Ohio Vocational Competency Assessment Program, 14, 56
Oklahoma, 42, 49, 82
Oklahoma Occupational Testing Center, 42-43
OOTC. *See* Oklahoma Occupational Testing Center
Outcome-oriented assessment, 4
  measures, 5
  national activities, 25
OVAE. *See* Office of Vocational and Adult Education
OVCAP. *See* Ohio Vocational Competency Assessment Program

Pennsylvania, 5, 78
Performance assessment
  and written testing, 58-59, 85
  critical issues, 9, 39
  definition, 7,9
  effects on instruction, 15, 58-59
  hands-on performance tests, 7, 10, 73, 75, 77-79, 90
  statewide systems of, 53, 64
Performance-based accountability
  changing themes, 32-34
  current requirements, 22-23
  evolution of legislation, 4-6
  federal regulations, 105-106
  legislative history, 26-32
  overview of 1990 amendments, 23-26
Performance standards, 2
  link to federal funding, 5, 25
  national system, 5- 6, 25-26
  purpose, 5, 51
  special populations, 5
  technical assistance to states, 104-105
Perkins Act of 1984, 4, 29-30, 51, 112
  1990 amendments, 1-2, 46-47, 83, 103, 112
    effects, 16
    overview, 23-24
  effectiveness, 22
  evolution, 29-30
  special populations, 30
  state responses, 13-17
Program improvement. *See* Program effectiveness

Reading for information test, 69
Research and demonstration, 107
Resource management test, 70
Reviews of local program effectiveness. *See* Effectiveness of vocational programs

Smith-Hughes Act of 1917, 4, 26-27, 111
  instructional hours per week, 4, 26
SOCAT. *See* Student Occupational Competency Achievement Testing
South Carolina, 56
Speaking test, 70
Special populations
  and performance standards, 3, 5, 105
  balanced approach, 30, 34
  federal priority, 22, 24, 25
Standardized testing, 16, 39. *See also* Academic skills
  definition, 9
  effectiveness, 59
State survey
  design and methodology, 37-38
State-local program reviews. *See* Effectiveness of vocational education
Student Occupational Competency Achievement Testing, 67
  administration, 79-80
  future, 80
  overview, 76
  use, 78
Student Occupational Competency Testing Institute, 10-11

Student profiles and portfolios, 7, 43-44

Teacher Occupational Competency Testing, 76, 77, 78
Teacher qualifications, 4
Teamwork test, 69
Technical skills. *See* Broad technical skills
Tennessee, 11, 82
Test administration and reporting, 70-71
Test of Adult Basic Skills, 37, 49
Testing. *See also specific tests by name*
  differences in state policies, 41
  glossary of terms, 8
  mandatory, 41-42
  policies, 2, 38-41
  resources, 2, 6-7, 10-11, 42-44
  state policies, 35-65
Texas, 56
TIERS. *See* Title I Evaluation and Reporting System
Title I Evaluation and Reporting System, 31-32
TOCT. *See* Teacher Occupational Competency Testing
Types of skills, 2, 3
  changes in skills assessed, 15-16
  mix of skills assessed, 59-61

U.S. Department of Education, 2
  industry skill standards projects, 18
  initiative to develop standards, 107-108, 109
  regulations and monitoring, 105-106
  role, 103-104
  technical assistance efforts, 104-105
U.S. Department of Labor, 31
  industry skill standards projects, 18, 91
  initiative to develop standards, 107-108

V-TECS. *See* Vocational-Technical Education Consortium of the States
Vocational-academic integration, 16-17, 44, 50, 63-64
Vocational aptitudes, 84-91, 85
  aptitude versus achievement, 85
  types of vocational aptitude tests, 17
Vocational Education Act of 1963, 4, 27-28, 111
  1968 amendments, 4, 28, 111
  1976 amendments, 4, 28, 111
Vocational skills, 63
  changes in skills assessed, 15, 63
  definition, 2, 3, 46
  grade levels of, 48
  integrating with academic skills, 63
  sources of skill information, 49
  state assessment policies, 11-12, 17, 38
Vocational-Technical Education Consortium of the States, 10-11, 41-42, 67
  development process for catalogs and instructional tools, 73-74
  future, 76
  item banks, 74-76
  origin and purpose, 72-73
Voluntary Skill Standards and Certification Program. *See* Industry skill standards

Wages
  effects of technical skills on, 90
West Virginia, 56
Work Keys, 10, 14, 56
  assessment development process, 70
  assessments, 69-70
  design, 68-69
  implementation, 72
  job profiling, 71-72
  origin, 67-68
Written testing, 7, 53
  components, 58
  critical issues, 9
  definition, 36
  effects on instruction, 85
  listening and writing tests, 69
  state assessment policies, 12, 15

Superintendent of Documents **Publications** Order Form

Order Processing Code:
*7339

P3

Telephone orders (202) 783-3238
(The best time to call is between 8-9 a.m. EST.)
To fax your orders (202) 512-2250
Charge your order. It's Easy!

☐ **YES**, please send me the following:

_____ copies of ***Testing and Assessment in Vocational Education (128 pages),***
S/N 052-003-01367-3 at $7.50 each.

The total cost of my order is $_____. International customers please add 25%. Prices include regular domestic postage and handling and are subject to change.

_____
(Company or Personal Name) (Please type or print)

_____
(Additional address/attention line)

_____
(Street address)

_____
(City, State, ZIP Code)

_____
(Daytime phone including area code)

_____
(Purchase Order No.)

Please Choose Method of Payment:

☐ Check Payable to the Superintendent of Documents
☐ GPO Deposit Account ☐☐☐☐☐☐☐ – ☐
☐ VISA or MasterCard Account
☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐
☐☐☐☐ (Credit card expiration date)

Thank you for your order!

_____
(Authorizing Signature) (03/94)

YES NO
May we make your name/address available to other mailers? ☐ ☐

Mail To: New Orders, Superintendent of Documents, P.O. Box 371954, Pittsburgh, PA 15250-7954

THIS FORM MAY BE PHOTOCOPIED

Superintendent of Documents **Publications** Order Form

Order Processing Code:
*7339

P3

Telephone orders (202) 783-3238
(The best time to call is between 8-9 a.m. EST.)
To fax your orders (202) 512-2250
Charge your order. It's Easy!

☐ **YES**, please send me the following:

_____ copies of ***Testing and Assessment in Vocational Education (128 pages),***
S/N 052-003-01367-3 at $7.50 each.

The total cost of my order is $_____. International customers please add 25%. Prices include regular domestic postage and handling and are subject to change.

_____
(Company or Personal Name) (Please type or print)

_____
(Additional address/attention line)

_____
(Street address)

_____
(City, State, ZIP Code)

_____
(Daytime phone including area code)

_____
(Purchase Order No.)

Please Choose Method of Payment:

☐ Check Payable to the Superintendent of Documents
☐ GPO Deposit Account ☐☐☐☐☐☐☐ – ☐
☐ VISA or MasterCard Account
☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐☐
☐☐☐☐ (Credit card expiration date)

Thank you for your order!

_____
(Authorizing Signature) (03/94)

YES NO
May we make your name/address available to other mailers? ☐ ☐

Mail To: New Orders, Superintendent of Documents, P.O. Box 371954, Pittsburgh, PA 15250-7954

THIS FORM MAY BE PHOTOCOPIED

Office of Technology Assessment

The Office of Technology Assessment (OTA) was created in 1972 as an analytical arm of Congress. OTA's basic function is to help legislative policymakers anticipate and plan for the consequences of technological changes and to examine the many ways, expected and unexpected, in which technology affects people's lives. The assessment of technology calls for exploration of the physical, biological, economic, social, and political impacts that can result from applications of scientific knowledge. OTA provides Congress with independent and timely information about the potential effects—both beneficial and harmful—of technological applications.

Requests for studies are made by chairmen and ranking minority members of standing committees of the House of Representatives or Senate, by the Technology Assessment Board, the governing body of OTA, or by the Director of OTA in consultation with the Board.

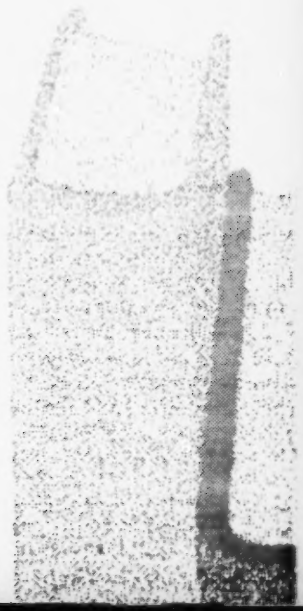
The Technology Assessment Board is composed of six members of the House, six members of the Senate, and the OTA Director, who is a nonvoting member.

OTA has studies under way in nine program areas: energy and materials; industry, technology, and employment; international security and commerce; biological and behavioral sciences; food and renewable resources; health; telecommunication and computing technologies; oceans and environment; and science, education, and transportation.

BEST COPY AVAILABLE

OFFICE OF
TECHNOLOGY ASSESSMENT

CONGRESS OF
THE UNITED STATES



OTA-BP-SET-123

MARCH 1994